National Library of Canada
Bibliothèque nationale du Canada

Canadian Theses Service
Service des thèses canadiennes

Ottawa, Canada
K1A 0N4

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Previously copyrighted materials (journal articles, published tests, etc.) are not filmed.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

Les documents qui font déjà l'objet d'un droit d'auteur (articles de revue, tests publiés, etc.) ne sont pas microfilmés.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30.

NL-339 (r. 88/04)

Canada

The University of Alberta

Matrix Padé Forms and Inverses of Block Hankel Matrices

by

© George Labahn

A thesis

submitted to the Faculty of Graduate Studies and Research

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Computing Science

Edmonton, Alberta

Fall, 1988

Permission has been granted to the National Library of Canada to microfilm this thesis and to lend or sell copies of the film.

The author (copyright owner) has reserved other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without his/her written permission.

L'autorisation a été accordée à la Bibliothèque nationale du Canada de microfilmer cette thèse et de prêter ou de vendre des exemplaires du film.

L'auteur (titulaire du droit d'auteur) se réserve les autres droits de publication; ni la thèse ni de longs extraits de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation écrite.

ISBN 0-315-45692-2

The University of Alberta

Release Form

Name of Author: George Labahn

Title of Thesis: *Matrix Padé Forms and Inverses of Block Hankel Matrices*

Degree for Which Thesis was Presented: Doctor of Philosophy

Year This Degree Granted: 1988

Permission is hereby granted to The University of Alberta Library to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only.

The author reserves other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without the author's written permission.

George Labahn

Permanent Address: 11704 - 44 Avenue

Edmonton, Alberta

Canada

Dated 30 June 1988

The University of Alberta

Faculty of Graduate Studies and Research

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research, for acceptance, a thesis entitled *Matrix Padé Forms and Inverses of Block Hankel Matrices* submitted by **George Labahn** in partial fulfillment of the requirements for the degree of Doctor of Philosophy

.....

Supervisor

.....

.....

.....

.....

Date June 17, 1988

Dedication

For the children

Philip Labahn, Eve Coppinger, and Allison Leslie

Abstract

For matrix power series with coefficients over a field, the notion of a matrix power series remainder sequence and its corresponding cofactor sequence are introduced and developed. An algorithm for constructing these sequences is presented.

It is shown that the cofactor sequence yields directly a sequence of Padé fractions for a matrix power series represented as a quotient $B(z)^{-1}A(z)$. When $B(z)^{-1}A(z)$ is normal, the complexity of the algorithm for computing a Padé fraction of type (m,n) is $O(p^3(m+n)^2)$, where p is the order of the matrices A(z) and B(z).

For power series which are abnormal, for a given (m,n), Padé fractions may not exist. However, it is shown that a generalized notion of Padé fraction, the Padé form, introduced in this paper does always exist and can be computed by the algorithm. In the abnormal case, the algorithm can reach a complexity of $O(p^3(m+n)^3)$, depending on the nature of the abnormalities. In the special case of a scalar power series, however, the algorithm complexity is $O((m+n)^2)$, even in the abnormal case.

Matrix Padé forms are also used to give a set of new formulae for the inverse of a block Hankel matrix. The formulae are practical because our algorithm allows for easy computation of these forms. In addition, a by-product of our formulae and the algorithm allows the computation of all inverses of nonsingular principal minors.

By using Frobenius-type identities between certain matrix Padé forms, the inversion formulae are shown to generalize the formulae of Gohberg-Heinig, and, in the scalar case, the formulae of Gohberg-Semencul and Gohberg-Krupnik.

The new formulae have the significant advantage of requiring only that the block Hankel matrix itself be nonsingular. The others require, in addition, that certain submatrices be nonsingular.

Acknowledgements

The author would like to take this opportunity to thank the people who helped make this work possible. As with any research project, numerous people were consulted at all stages of its development. Their time and often insightful criticism of this work sparked new ideas and made old ones clear. In particular, I would like to thank the following people.

My first acknowledgements must go to my supervisor Stan Cabay. During the three years I spent under his supervision I grew both academically and professionally. May everyone have such a helpful supervisor.

Second, I would like to thank the members of my examining committee for their time spent in reviewing this thesis.

Without the resources and technical support provided by The Department of Computing Science at the University of Alberta this thesis would not have been possible. In addition a special thanks goes to the Department of Mathematics at the University of Alberta for their support and kindness over the years.

Also, none of this would have been possible without the unlimited patience and kindness of my wife, Laura Leslie, and the ever present encouragement of my parents.

Table of Contents

Chapter	Page
Chapter 1: Introduction	1
Chapter 2: Matrix Padé Forms	12
2.1. Introduction	12
2.2. Basic Definitions	13
2.3. Matrix Padé Fractions.	17
Chapter 3: The MPADE Algorithm	22
3.1. Matrix Power Series Remainder Sequences	22
3.2. The Algorithm	31
3.3. Complexity of the MPADE Algorithm	36
Chapter 4: Inverse Formulae of Block Hankel Matrices	39
4.1. Introduction	39
4.2. Nonsingular Block Hankel Matrices and Matrix Padé Forms	40
4.3. The Off-diagonal Inverse Formulae	46
4.4. The Anti-diagonal Inverse Formulae	49
Chapter 5: Other Inverse Formulae	54
5.1. Introduction	54
5.2. The Inverse Formulae of Gohberg-Heinig	55
5.3. The Inverse Formulae of Gohberg-Krupnik	60
5.4. Additional Inverse Formulae	68
5.5. Computation of Block Hankel Inverses	68

Chapter 6: Summary and Suggestions for Future Research 71

Chapter 7: Appendix: Validity of Initialization in MPADE Algorithm 76

References 78

Chapter 1

Introduction

Let

$$A(z) = \sum_{i=0}^{\infty} a_i z^i, \tag{1.1}$$

where a_i, i = 0, ... , is a p x p matrix with coefficients from a field K, be a formal power series. Loosely speaking, a matrix Padé approximant of A(z) is an expression of the form $U(z) \cdot V(z)^{-1}$, or $V(z)^{-1} \cdot U(z)$, where U(z) and V(z) are matrix polynomials of degree at most m and n, respectively, whose expansion agrees with A(z) up to and including the term z^{m+n}. Mathematically this is equivalent to the expressions

$$A(z)V(z) - U(z) \quad or \quad V(z)A(z) - U(z) \tag{1.2}$$

being zero for the first m+n+1 terms.

Padé approximants have many applications in mathematics and in engineering-related disciplines. Applications include numerical computations for special power series such as the Gamma function (c.f., Nemeth and Zimanyi [37]), algorithms in the field of numerical analysis (c.f., Gragg [26]), triangulation of block Hankel and Toeplitz matrices (c.f., Rissanen [41]), solution of linear systems of equations with Hankel or Toeplitz coefficient matrices (c.f., Rissanen [39]), digital filtering theory (c.f., Bultheel [12] and Brophy and Salazar [9]) and linear control theory (c.f., Elgerd [22]).

As an illustration we can single out three areas where formal power series and related rational expressions come into play.

Example 1.1:

Consider the initial valued nonlinear differential equation given by

$$y \cdot \ddot{y} + \dot{y}^2 = \frac{3}{(1+2z)^3}, \quad y(0) = 1, \quad \dot{y}(0) = -3/4. \tag{1.3}$$

If we make the assumption that y is analytic in z in a neighborhood of 0, then the coefficients of the

MacLaurin expansion can be determined as follows. The first two coefficients are obtained from the initial conditions while the third is determined by substituting 0 into the differential equation and solving for $\ddot{y}(0)$. This gives

$$y = 1 - 3/4z + 39/32z^2 + \cdots . \tag{1.4}$$

Higher order coefficients can be obtained by differentiating the equation, substituting 0, and solving for $y^{(n)}(0)$. This process is only useful as an approximation since one must stop after a finite number of steps. The result is an approximation of y by a polynomial. However, polynomials are often poor approximations, especially for large z. Instead, one can use a Padé approximation, which, when truncated at the same position, produces the same polynomial. For example, the rational expression

$$\frac{1 + 7/8z}{1 + 13/8z} \tag{1.5}$$

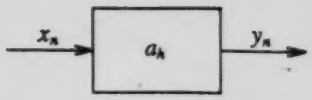
also has the same first three terms as the power series expansion for y (this is the Padé fraction of type (1,1) for y). This rational expression results in a much closer approximation to the real answer

$$\sqrt{\frac{2+z}{2+4z}} \tag{1.6}$$

especially for large z. There is, for example, only an 8% difference at infinity (c.f., Baker and Graves-Morris [3]). In this case Padé approximation is used as a method of improving an existing approximation.

Example 1.2:

Consider a linear system having discretized input x_n and discretized output y_n. The outputs are determined by convolution of the inputs with a unit impulse response a_k. (c.f., Bultheel [12])



If we think of these discrete sequences in terms of their z-transforms, i.e., as power series

$$X(z) = \sum x_n z^n, \quad Y(z) = \sum y_n z^n, \quad A(z) = \sum a_n z^n, \tag{1.7}$$

then convolution of the impulse response function goes over to polynomial multiplication which results in

$$Y(z) = A(z) \cdot X(z). \tag{1.8}$$

A(z) is called the transfer function of the system. The coefficents of the transfer function are usually determined theoretically or experimentally by measuring the output of a signal from an input that isolates on the particular response. An approximation of A(z) by a rational expression

$$\frac{U(z)}{V(z)} = \frac{\sum_{i=0}^{m} u_i z^i}{\sum_{i=0}^{n} v_i z^i}, \quad \text{with} \quad v_0 = 1, \tag{1.9}$$

combined with (1.8) gives

$$V(z) \cdot Y(z) = U(z) \cdot X(z). \tag{1.10}$$

Thus

$$y_t = \left[u_0 x_t + \cdots + u_m x_{t-m} \right] - \left[v_1 y_{t-1} + \cdots + v_n y_{t-n} \right] \tag{1.11}$$

which is a system where the output at time t depends on a particular combination of the last m+1 inputs and the previous n outputs. In this case rational approximation is used to obtain a feedforward/feedback model for a linear system (c.f., Gragg and Lindquist [27]).

A variation of this example is to find a closest model to the system having a particular type of rational expression, for example one of the form

$$\frac{u_0}{v_0 + v_1 z + \cdots v_n z^n} . \tag{1.12}$$

By closest, we mean in the least squares sense. Solving the normal equations of this over-determined system results in a Yule-Walker system of equations (c.f, chapter 4) where the coefficients of the system are the autocorrelation coefficients of the sequence a_k (c.f., Wakita [46] or Makhoul [34]). This in turn can be interpreted as a Padé approximation of a particular type for the autocorrelation power series.

Example 1.3

For the power series

$$\frac{1 - 4z^2}{1 + 3z - 3z^2} = 1 - 1/3z + O(z^3), \tag{1.11}$$

the Padé fraction of type (0,2) is given by

$$\frac{1}{1 + 3z + z^2} \tag{1.12}$$

since

$$\frac{1}{1 + 3z + z^2} = \frac{1 - 4z^2}{1 + 3z - 3z^2} + O(z^3). \tag{1.13}$$

Multiplying denominators gives

$$1 + 3z - 3z^2 = (1 + 3z + z^2)(1 - 4z^2) + O(z^3), \tag{1.14}$$

or, more precisely,

$$1 + 3z - 3z^2 = (1 + 3z + z^2)(1 - 4z^2) + z^3(12 + 4z). \tag{1.15}$$

Multiplying both sides by z^{-4} and substituting z for z^{-1} (this is the same as reversing the order of the coefficients of the polynomials) results in

$$z^4 + 3z^3 - 3z^2 = (z^2 + 3z + 1)(z^2 - 4) + (12z + 4) \tag{1.16}$$

which is just polynomial division with remainder. This process shows that polynomial division is equivalent to Padé approximation (of a specific type). More generally, algorithms that are based on polynomial division can also be interpreted in terms of Padé approximation. For example, for any two polynomials A(z) and B(z), the extended Euclidean GCD algorithm calculates a sequence of expressions of the form

$$A(z)S_i(z) + B(z)T_i(z) = R_i(z), \tag{1.17}$$

satisfying various algebraic properties (c.f., [35]), with the last remainder $R_{last}(z)$ being the GCD of A(z) and B(z). If m_i and n_i are the degrees of $S_i(z)$ and $T_i(z)$, respectively, then reversing coefficients in equation (1.16) results in

$$A^{(r)}(z)S_i^{(r)}(z) + B^{(r)}(z)T_i^{(r)}(z) = z^{m_i+n_i+1}R_i^{(r)}(z). \tag{1.18}$$

where the superscript (r) denotes a polynomial having coefficients in reversed order. Thus the extended Euclidean GCD algorithm can be viewed as a sequence of Padé approximations for the quotient power series $B^{(r)}(z)^{-1}A^{(r)}(z)$.

Note that, although the examples are presented in the scalar case, these are equally valid in the multidimensional case. Matrix Padé approximants can be used when we have a system of nonlinear differential equations with initial conditions. Multidimensional rational approximants are used with linear systems having more than one input and one output.

These examples also illustrate problems that must be considered when generalizing from the scalar case to the matrix case. In example 1.2, commutativity becomes an issue, since equation (1.9) requires that the denominator divide on the left. A denominater dividing on the right will not give a desired model because of the lack of commutativity. Example 1.3 points out the problem of singularity in the matrix case. Matrix polynomial division is equivalent to specific types of matrix Padé approximation, just as in the

scalar case. However, there is no Euclidean algorithm in the case of matrix polynomials, because nonzero coefficents may be singular, and hence one cannot always divide one nonzero coefficient into another nonzero coefficient. For more on the problem of generalizing scalar polynomials to their matrix counterparts see for example Dennis et al [19] .

An added consideration when generalizing rational approximants is that, unlike the scalar case, the definition of a matrix Padé approximant can be made formal in a variety of ways. For example, the approximation (1.2) can be written as

$$A(z) \simeq U'(z) \cdot V'(z)^{-1} = \left\{ U'(z) \cdot adj(V'(z)) \right\} \cdot (det(V'(z))^{-1} = \left\{ U(z) \right\} \cdot V(z)^{-1} \qquad (1.19)$$

so that V(z) is a scalar polynomial while U(z) can be a p x q matrix. This is the approach taken by Rissanen [40] . Typically, however, U(z) and V(z) are p x p polynomial matrices, and V(z) is further restricted by the condition that the constant term, V(0), is invertible. In this thesis, we call such approximants matrix Padé fractions, which is consistent with the scalar (p=1) case.

For a particular m and n, however, matrix Padé fractions need not exist. Therefore, in this thesis we introduce the notion of a matrix Padé form, in which the condition of invertibility of V(0) is relaxed. The definition is a generalization of a similar one given in the scalar case. It is shown that matrix Padé forms always exist, but that they may not be unique. In general, matrix Padé forms need not have an invertible denominator, V(z). However, for m and n given, by obtaining a basis for all the Padé forms, we are always able to construct a matrix Padé form with an invertible V(z), in the case that one does indeed exist. This, along with additional properties of matrix Padé forms, provides the contents of chapter 2.

Chapter 3 is concerned with the problem of calculating matrix Padé approximants. In the scalar case, existing algorithms can be classified into two categories, those that require that A(z) be a normal power series (c.f., Gragg [26]) and those that do not require this restriction. Algorithms that calculate Padé approximants for normal power series include the ε-algorithm of Wynn [48] , the η-algorithm of Bauer [5] , the Q-D algorithm of Rutishauser [42] along with the well known algorithms of Trench [45] and Levinson

[33] Algorithms that are successful in the degenerate non-normal case can further be divided into two subcategories according to their complexity costs. The algorithms of Bultheel [11] , Rissanen [39] , McEliece and Shearer [35] , and the slow algorithm of Cabay and Choi [16] calculate the Padé approximant of type (m,n) with a complexity of $O((m+n)^2)$ operations. The algorithms of Brent et al [8] , Sugiyama [44] and the fast algorithm of Cabay and Choi [16] have a complexity of only $O((m+n)\log^2(m+n))$. See the article by Sugiyama for a table listing well known algorithms including their restrictions along with complexities.

Most scalar algorithms can be generalized if the matrix power series is assumed to be normal. While there are examples where the lack of commutativity is a concern (c.f., Draux [20]) this is not a problem with most algorithms. Examples of such generalizations to the matrix case include those given by Bose and Basu [7] , Bultheel [10] , Rissanen [41] , and Starkand [43] . The fast methods of Brent et al [8] , Cabay and Choi [16] , and Sugiyama [44] also work over matrix coefficients, under these strict normality conditions. Fast polynomial multiplication using fast Fourier transforms is required to obtain the improved complexity for these fast algorithms. Since matrix polynomials can be considered as matrices having polynomial coefficients, multiplying two matrix polynomials is the same as a series of scalar polynomial multiplications. Thus fast techniques carry over to the matrix case.

All these algorithms fail, however, if the normality condition is dropped. Without normality there arise problems with matrix singularity. That is, in each algorithm specific nonzero coefficients need to be invertible, a requirement that cannot be guaranteed in the case of matrix coefficients. These specific nonzero coefficients are nonsingular when normality is assumed. An algorithm that calculates Padé approximants in a non-normal case is given by Labahn [32] . However, in this algorithm there are still strict conditions (near-normality) that need to be satisfied by the power series before Padé approximants can be calculated. Some other matrix algorithms that require normality will also work in this slightly less restrictive case (e.g., Rissanen's algorithm).

In chapter 3 we present an algorithm, MPADE, for computing matrix Padé forms for a given matrix power series. The algorithm overcomes the problem of singularity and so, unlike other algorithms, there

are no restrictions placed on the power series in order that MPADE succeed. Central to the development of MPADE are the notions of a matrix power series remainder sequence and the corresponding cofactor sequence, which are introduced in this chapter. These are generalizations of notions developed by Cabay and Kossowski [17] for power series over an integral domain. The cofactor sequence computed by MPADE yields a sequence of matrix Padé fractions along a specific off-diagonal path of the Padé table for A(z).

A complexity analysis of the algorithm is given. For normal power series, we show that the complexity of MPADE is $O(p^3 \cdot (m+n)^2)$ operations in K. This is the same complexity as some of the algorithms proposed by Bultheel [10] , Bose and Basu [7] , Starkand [43] , and Rissanen [41] . Overcoming the problem of singularity introduces a degree of uncertainty in the cost complexity, however. In the abnormal case, the algorithm is usually of complexity $O(p^3 \cdot (m+n)^2)$, but can reach $O(p^3 \cdot (m+n)^3)$ operations in K in some pathological cases. In the scalar case, however, MPADE always has a complexity of $O((m+n)^2)$, regardless of the degree of abnormality.

The concept of a matrix Padé approximant is closely related to matrices having a block Hankel structure. These are matrices of the form

$$H_{m,n} = \begin{bmatrix} a_{m-n+1} & \cdots & a_m \\ . & & . \\ . & & . \\ . & & . \\ a_m & \cdots & a_{m+n-1} \end{bmatrix} \tag{1.19}$$

where the a_i are p x p matrices. The connection between the two follows from equation (1.2) which allows the denominator to be calculated as solutions of a system of equations having as their coefficient matrix a block Hankel matrix. Systems of equations having as their coefficient matrix a Hankel matrix are encountered in such applications as digital filtering, linear control theory and image processing.

Chapter 4 deals with the problem of inverting block Hankel matrices. The approach uses the relationship between these matrices and a related set of matrix Padé forms to determine both when a Hankel

matrix is nonsingular and also how to calculate its inverse. We show that necessary and sufficient conditions for a block Hankel matrix to be nonsingular is the existence of four matrix Padé forms satisfying specific conditions. In addition, a set of closed formulae for the inverse is given in terms of these four matrix Padé forms.

The existence of closed formulae for Hankel matrices is not new. In the scalar case well known formulae of Gohberg and Semencul [23] gives $H_{m,n}^{-1}$ in terms of only the first and last columns of the inverse. Gohberg and Krupnik [24] give a formula for the inverse in terms of only the last two columns of $H_{m,n}^{-1}$.

In the nonscalar case, when $p > 1$, additional problems are encountered in obtaining a closed formula for the inverse. A set of closed formulae of Gohberg and Heinig [25] , give the inverse in terms of the first and last columns together with the first and last rows of the inverse.

All of the above formulae depend on the ability to perform certain bordering operations that lend themselves well to matrices with a Hankel structure. However, these bordering operations require the imposition of certain additional restrictions on $H_{m,n}$. For both the Gohberg-Semencul and Gohberg-Heinig formulae, for example, the submatrix $H_{m,n-1}$ must also be nonsingular; whereas, for the Gohberg-Krupnik formula the submatrix $H_{m-1,n-1}$ must also be nonsingular. Inverse formulae are then also given for the relevant submatrices.

By using matrix Padé forms, we are able to avoid bordering techniques and so require only that $H_{m,n}$ be nonsingular. In addition we show, in chapter 5, that formulae that are based on bordering techniques can easily be determined from our formulae by taking advantage of relationships (e.g., Frobenius identities) that exist for neighbors in a Padé table. Thus, by using Padé forms we unify the work done previously in this field.

A major advantage of a closed inverse formula is that it allows for efficient algorithms to calculate the inverses of Hankel matrices. This efficiency comes both in the cost complexity of calculating the inverse and also in the amount of storage required for the final result.

As an application of both our inverse formulae and the MPADE algorithm, we obtain an algorithm for calculating $H_{m,n}^{-1}$. The resulting algorithm has many advantages for our situation. It is successful without any preconditions placed on the original power series. As a by-product, we obtain inverses for all the principal minors of $H_{m,n}$ that are nonsingular. Also, it is iterative on n, allowing cost savings in implementation. The complexity of the resulting algorithm for an n x n matrix is generically $O(p^3n^2)$, although there are pathological cases where it can be as high as $O(p^3n^3)$ (for example, when all the principal minors of $H_{m,n}$ are singular). This compares with other nonscalar methods (c.f., Akaike [1], Rissanen [41], Bose and Basu [7]) which are also of complexity $O(p^3n^2)$, but which succeed only when all principal minors are nonsingular. In the scalar case, however, the cost complexity of MPADE is $O(n^2)$, regardless of the types of singularities found in $H_{m,n}$. This compares favorably with the method described by Rissanen [39], which is of complexity $O(n^2)$ and succeeds in the degenerate case. The $O(n^2)$ methods of Trench [45], Zohar [49], and Kailath et al [30], on the other hand, fail whenever a principal minor of $H_{m,n}$ is singular.

When fast polynomial multiplication methods are available, in the scalar case, the required Padé forms can be calculated by the EMGCD algorithm of Brent et al [8] or the off-diagonal algorithm of Cabay and Choi [16] with a complexity of $O(n \log^2 n)$. The algorithm is also iterative on n and produces the inverses of some of the nonsingular principal minors as a by-product. The $O(n \log^2 n)$ methods of Bitmead and Anderson [6] and de Hoog [28], on the other hand, succeed only in the nondegenerate case. We note that the last two algorithms all use fast $O(n \log^2 n)$ algorithms to determine both the first and last columns of the inverse, and then use one of the inverse formula of Gohberg and Semencul to obtain the inverse. Many of the inefficiencies or restrictions are due to the limitations of this inverse formula.

In the nonscalar case, fast algorithms can also be used to calculate the required Padé forms, but under some restrictions. If the block matrix is positive definite (or, more generally, if the associated power series is nearly-normal (c.f., chapter 3)), for example, and fast polynomial multiplication is allowed, then the inverse formulae can be calculated using the fast algorithm of Labahn [32] with complexity $O(p^3 \cdot n \log^2 n)$. This algorithm is also iterative and calculates the inverses of some of the nonsingular principal minors as a

by-product. The algorithm of Bitmead and Anderson, generalized to the nonscalar case by the use of the inverse formulae of Gohberg and Heinig, is also of complexity $O(p^3 \cdot n \log^2 n)$, but works only in the normal case.

Chapter 2

Matrix Padé Forms

2.1. Introduction

Let

$$A(z) = \sum_{i=0}^{\infty} a_i z^i, \tag{2.1}$$

where a_i, i = 0, ... , is a p x p matrix with coefficients from a field K, be a formal power series. In this chapter we present the basic definitions and properties of a matrix Padé approximant for A(z). Roughly speaking this is an expression of the form $U(z) \cdot V(z)^{-1}$, or $V(z)^{-1} \cdot U(z)$, where U(z) and V(z) are p x p matrix polynomials of degree at most m and n, respectively, whose expansion agrees with A(z) up to and including the term z^{m+n}.

Typically V(z) is further restricted by the condition that the constant term, V(0), is invertible (c.f., Bose and Basu [7] , Bultheel [10] , and Starkand [43]). In this thesis, we call such approximants matrix Padé fractions, which is consistent with the scalar (p=1) case (c.f., Gragg [26]).

For a particular m and n, however, matrix Padé fractions need not exist. Therefore we introduce the notion of a matrix Padé form, in which the condition of invertibility of V(0) is relaxed. The definition is a generalization of a similar one given for the scalar case (c.f., Gragg [26]). We show that matrix Padé forms always exist, but that they may not be unique. In general, matrix Padé forms need not have an invertible denominator, V(z). However, for m and n given, by obtaining a basis for all the Padé forms, we are also able to construct a matrix Padé form with an invertible V(z), in the case that one does exist. In addition, properties that will be used (in chapter 3) to establish an algorithm to construct such forms are also included in this chapter.

For purposes of presentation, we adopt the following notation. We let D denote the non-commutative ring of p x p matrices over a field. The domain of formal power series with coefficients over D and indeter-

minant z is denoted by D[[z]]. For any $A(z) \varepsilon D[[z]]$, A(z) is formally represented by

$$A(z) = \sum_{i=0}^{\infty} a_i z^i, \tag{2.2}$$

where the coefficients $a_i \varepsilon D$ are always written in lower case. The domain of polynomials (finite power series) over D with indeterminant z is denoted by D[z]. Any polynomial $P_n(z) \varepsilon D[z]$ is represented formally by

$$P_n(z) = \sum_{i=0}^{n} p_i z^i, \tag{2.3}$$

where again the coefficients $p_i \varepsilon D$ are written in lower case. The degree of $P_n(z)$ (i.e., the largest i such that $p_i \neq 0$) is denoted by $\partial(P_n(z))$.

2.2. Basic Definitions

Let

$$A(z) = \sum_{i=0}^{\infty} a_i z^i \ , \ B(z) = \sum_{i=0}^{\infty} b_i z^i \ \varepsilon \ D[[z]] \tag{2.4}$$

be formal power series with coefficients from the ring D of p x p matrices over some field K. Throughout this paper it is assumed that the leading coefficient, b_0, of B(z) is an invertible matrix. For non-negative integers m and n, let

$$U(z) = \sum_{i=0}^{m} u_i z^i \ , \ V(z) = \sum_{i=0}^{n} v_i z^i \ \varepsilon \ D[z] \tag{2.5}$$

denote p x p matrix polynomials.

Definition 2.1: The triple (U (z),V (z),W(z)) is defined to be a **Right Matrix Padé Form** (RMPFo) of type (m,n) for the pair (A(z),B(z)) if

I. $\partial(U(z)) \leq m$, $\partial(V(z)) \leq n$,

II. $A(z)\cdot V(z) + B(z)\cdot U(z) = z^{m+n+1}W(z)$ *with* $W(z) \in D[[z]]$ and . (2.6)

III. The columns of $V(z)$ are linearly independent over the field K.

The matrix polynomials U(z), V(z), and W(z) are usually called the **right numerator, denominator,** and **residual** (all of type (m,n)), respectively. Note that when B(z) = -I, Definition 2.1 corresponds to the definition of Padé form for a single matrix power series A(z) given in Labahn [32] .

There is an equivalent definition for a left matrix Padé form (LMPFo). Condition II is replaced with an equivalent version with matrix multiplication by U(z) and V(z) being on the left. Condition III is replaced with the condition that the rows, rather than the columns, of the denominator are linearly independent over the base field K.

However, there is a one-to-one correspondence between RMPFo's and LMPFo's. By taking the transposes of the matrices on both sides of equation (2.6), it follows that

$$V^t(z)\cdot A^t(z) + U^t(z)\cdot B^t(z) = z^{m+n+1}W^t(z)\,. \tag{2.7}$$

The degree and order conditions are identical. It is clear that if (U(z),V(z),W(z)) is a RMPFo for (A(z), B(z)), then $(U^t(z), V^t(z), W^t(z))$ is a LMPFo for $(A^t(z), B^t(z))$. Thus, any algorithm that calculates a right matrix Padé form of a certain type can also be used to calculate the left matrix Padé form of the same type.

For ease of discussion, we use the following notation. For any matrix polynomial

$$U(z) = u_0 + u_1 z + \cdots + u_k z^k \in D[z], \tag{2.8}$$

we write U (i.e., the same symbol but without the z variable) to mean the p(k+1) by p vector of matrix coefficients

$$U = \left[u_0, u_1, \cdots, u_k \right]^t. \tag{2.9}$$

Let

$$S_{m,n} = \begin{bmatrix} a_0 & & & b_0 & & \\ & \ddots & & & \ddots & \\ & & a_0 & & & b_0 \\ a_{m+n} & \cdots & a_m & b_{m+n} & \cdots & b_n \end{bmatrix} \qquad (2.10)$$

denote a Sylvestor matrix for A(z) and B(z) of type (m,n). Then equation (2.6) can be written as

$$S_{m,n} \cdot \begin{bmatrix} V \\ U \end{bmatrix} = 0. \qquad (2.11)$$

Theorem 2.2: (Existence of Matrix Padé Forms) For any pair of power series (A(z), B(z)) and any pair of nonzero integers (m,n), there exists a RMPFo of type (m,n).

Proof: Let X denote a vector of length p(m+n+2), and consider the homogeneous system of linear equations

$$S_{m,n} \cdot X = 0. \qquad (2.12)$$

Because $S_{m,n}$ has p(m+n+1) rows, it follows that (2.12) has at least p linearly independent solutions. Let $\left[V,U\right]^t$ denote p such solutions arranged by columns. Then $\left[V,U\right]^t$ satisfies (2.11), and consequently U(z) and V(z) determined according to the convention (2.8) and (2.9) satisfy (2.6). Clearly, the pair (U(z),V(z)) also satisfies condition I in Definition 2.1. The linear independence of the columns of V(z) is equivalent to the linear independence of the columns of $\left[V,U\right]^t$ because of the nonsingularity of b_0. Hence condition III of Definition 2.1 also holds.

From the proof of Theorem 2.2, it follows that if $S_{m,n}$ has maximal rank, then Padé forms are unique up to multiplication of U(z) and V(z) on the right by a nonsingular matrix. On the other hand, if the rank of

$S_{m,n}$ is less than maximal, then more than one independent Padé form exists.

Example 2.3: Let B(z) = -I and

$$A(z) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} z^2 + \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} z^4 + \begin{bmatrix} -1 & 0 \\ -1 & 0 \end{bmatrix} z^5 + \cdots . \tag{2.13}$$

With m = 2 and n = 3, a basis for the solution space of (2.9) is given by the two vectors

$$X_1 = \left[0, 1, 0, 0, 0, -1, 0, 0, 0, 1, 0, 0, 0, 0 \right]^t \tag{2.14}$$

and

$$X_2 = \left[0, 0, 0, 1, 0, 0, 0, -1, 0, 0, 0, 1, 0, 0 \right]^t . \tag{2.15}$$

Thus,

$$V = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & -1 \end{bmatrix}^t \tag{2.16}$$

and

$$U = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}^t \tag{2.17}$$

is a solution of (2.11), and the pair (U(z),V(z)), where

$$V(z) = \begin{bmatrix} 0 & 0 \\ 1-z^2 & z-z^3 \end{bmatrix} \quad \text{and} \quad U(z) = \begin{bmatrix} 0 & 0 \\ 1 & z \end{bmatrix} \tag{2.18}$$

defines a Padé form of type (2,3) for (A(z), B(z)).

In Example 2.3, note that that columns of V(z) are linearly independent over the field K, but that they are linearly dependent over the ring of polynomials $K[z]$ (i.e., V(z) is singular). Indeed, for this example, a RMPFo (U(z), V(z), W(z)) of type (2,3) for which V(z) is nonsingular cannot be found. The problem occurs because, although the solution space has dimension 2 when considered as a vector space over the

field K, it has only dimension 1 when considered as a module over the ring $K[z]$.

We note that having an invertible denominator is highly desirable, since usually the purpose of Padé forms is to approximate the infinite power series

$$-(B(z))^{-1} \cdot A(z) \tag{2.19}$$

by the finite rational form

$$U(z) \cdot (V(z))^{-1}, \tag{2.20}$$

where the approximation is to be exact for the first m+n+1 terms. When the denominator is singular, we cannot form this rational expression and this limits the usefulness of Padé approximation. For example, a singular denominator gives no information about the poles since every point is a pole in this case.

2.3. Matrix Padé Fractions.

One case when the denominator of a RMPFo is invertible is given by

Definition 2.4. The triple (U(z),V(z),W(z)) is said to be a **Right Matrix Padé Fraction** (RMPFr) of type (m,n) for the pair (A(z),B(z)) if

I. (U(z),V(z), W(z)) is a RMPFo of type (m,n) for (A(z),B(z)), and

II. The constant term, V(0), of the denominator is an invertible matrix.

Condition II ensures that the denominator, V(z), is an invertible matrix polynomial.

As in the case of Padé forms, there is an equivalent definition for a left matrix Padé fraction (LMPFr). Also, there is a correspondence between right matrix Padé fractions for (A(z),B(z)) and left matrix Padé fractions for $(A(z)^t, B(z)^t)$. It is interesting to note that a power series may have a matrix

Padé fraction on one side but not on the other. In Example 2.3, the power series A(z) does not have a right matrix Padé fraction of type (2,3), but it does have a left matrix Padé fraction of type (2,3). When a power series does have both a right and a left matrix Padé fraction of the same type, then the two resulting rational forms are equal (c.f., Baker [2]).

The problem with Padé fractions is that, as mentioned in the previous section, they do not always exist. However, let

$$T_{m,n} = \left[\begin{array}{ccc|ccc} a_0 & & & b_0 & & \\ \cdot & \cdot & & \cdot & \cdot & \\ \cdot & \cdot & & & & b_0 \\ \cdot & & a_0 & \cdot & & \cdot \\ & & \cdot & & & \cdot \\ & & \cdot & & & \\ a_{m+n-1} & \cdots & a_m & b_{m+n-1} & \cdots & b_n \end{array}\right] \tag{2.21}$$

and define

$$d_{m,n} = \begin{cases} 1, & m = 0\,,\, n = 0, \\ det\,(T_{m,n}), & otherwise. \end{cases} \tag{2.22}$$

Then, a sufficient condition for the existence of a RMPFr is given by

Theorem 2.5. If $d_{m,n} \neq 0$, then every RMPFo of type (m,n) is an RMPFr of type (m,n). In addition, a RMPFr of type (m,n) is unique up to multiplication on the right by a nonsingular p x p matrix having coefficients from the field K.

Proof: Equation (2.11) may be written

$$\begin{bmatrix} 0 & & b_0 & & \\ a_0 & & & & \\ & \ddots & & & \\ & a_0 & & b_0 & \\ & & \ddots & & \\ & & & & \\ a_{m+n-1} & \cdots & a_m & b_{m+n} \cdots & b_n \end{bmatrix} \begin{bmatrix} v_1 \\ \\ v_n \\ u_0 \\ \\ u_m \end{bmatrix} = - \begin{bmatrix} a_0 \\ \\ \vdots \\ \\ a_{m+n} \end{bmatrix} \cdot v_0 . \tag{2.23}$$

The matrix on the left of (2.23) is nonsingular, since $d_{m,n} \neq 0$ and b_0 is nonsingular. Thus, all the solutions of (2.23) can be obtained by assigning v_0 arbitrarily and solving (2.23) for the remaining components (v_1, ... v_n, u_0, ... , u_m). If v_0 is chosen to be a singular matrix, then the solution obtained by solving (2.23) violates condition III in the definition of Padé form. Thus, in this case, all RMPFo's are RMPFr's.

To show uniqueness, suppose (U(z),V(z),W(z)) and (U'(z),V'(z),W'(z)) are two RMPFr's for (A(z),B(z)). Then, v_0 and v'_0 are both nonsingular matrices with coefficients from the field K. Thus, there exists a nonsingular matrix, M, with coefficients from K satisfying $v_0 = v'_0 \cdot M$. It follows that

$$V(z) = V'(z) \cdot M \quad \text{and} \quad U(z) = U'(z) \cdot M \ , \tag{2.24}$$

while from (2.6) and (2.24)

$$z^{m+n+1} W(z) = z^{m+n+1} W'(z) \cdot M \tag{2.25}$$

and so uniqueness holds.

A basic requirement for results in both chapter 3 and chapter 4 is

Theorem 2.6. Let A(z) and B(z) be given by (2.4). If m and n are positive integers such that $d_{m,n} \neq 0$, then RMPFo's (P(z),Q(z),R(z)) of type (m-1,n-1) for (A(z),B(z)) are unique up to multiplication of P(z), Q(z) and R(z) on the right by a nonsingular matrix from K. In addition, the leading term R(0) of the residual is a

nonsingular matrix.

Proof: $S_{(m-1),(n-1)}$ can be obtained from $T_{m,n}$ by deleting the last block row (i.e., the last p rows). Since $T_{m,n}$ is of maximal rank p(m+n), then $S_{(m-1),(n-1)}$ has rank p(m+n-1). Consequently, the dimension of the solution space to

$$S'_{(m-1),(n-1)} \cdot X = 0 \tag{2.26}$$

is exactly p. Then, $[\,Q\,,P\,]^t$ is obtained by collecting by columns a basis for the solution space of (2.26). Clearly, if $[\,Q'\,;P'\,]^t$ and $[\,Q\,,P\,]^t$ are two such collections, then there exists a nonsingular matrix M from K such that

$$[\,Q\,,P\,]^t = [\,Q'\,,P'\,]^t \cdot M\,. \tag{2.27}$$

Thus, $P(z) = P'(z) \cdot M$ and $Q(z) = Q'(z) \cdot M$, and, from (2.6), $R(z) = R'(z) \cdot M$, proving uniqueness.

To prove the invertibility of R(0), let $r_0 = R(0)$ and suppose that r_0 is a singular p x p matrix. Then, there is a nonzero p x 1 vector X that satisfies

$$r_0 \cdot X = 0. \tag{2.28}$$

But, from (2.6) and (2.26), it follows that

$$T_{m,n} \cdot \begin{bmatrix} Q \\ P \end{bmatrix} = \begin{bmatrix} 0 \\ \cdot \\ 0 \\ r_0 \end{bmatrix}. \tag{2.29}$$

Thus,

$$T_{m,n} \cdot \begin{bmatrix} Q \\ P \end{bmatrix} \cdot X = 0\,. \tag{2.30}$$

Since the coefficient matrix for the above system is invertible, we deduce that

$$\begin{bmatrix} Q \\ P \end{bmatrix} \cdot X = 0\,. \tag{2.31}$$

But this contradicts the fact that the columns of $[\,Q\,,P\,]^t$ are made up of linearly independent vectors. This implies that r_0 is invertible.

The fact that Padé forms of type (m,n) and (m-1,n-1) are uniquely determined after suitable normalizations, when $T_{m,n}$ is nonsingular, allows us to prove such properties as argument invariance (c.f., Baker [2]) for the Padé forms computed by the algorithm MPADE given in chapter 3, and is central to our commutativity results of chapter 4.

Chapter 3

The MPADE Algorithm

3.1. Matrix Power Series Remainder Sequences

We define a **Right Matrix Padé Table** for (A(z), B(z)) to be any infinite two-dimensional collection of RMPFo's of type (m,n) for (A(z),B(z)) with m = 0, 1, ... and n = 0, 1, ... It is assumed that there is precisely one entry (i.e., one RMPFo) assigned to each position in the table. From Theorem 2.2, it follows that a right matrix Padé table exists for any given (A(z), B(z)). However, the table is not unique, because RMPFo's are not unique. This is unlike the definition of a Padé table for scalar power series (c.f. Gragg [26]), since here a Padé table consists of a collection of Padé fractions, which are unique.

A matrix power series pair (A(z),B(z)) is said to be **normal** (c.f., Bultheel [10]) if $d_{m,n} \neq 0$ for all m,n. For normal power series, it follows from Theorem 2.5 that every entry in the right matrix Padé table is a RMPFr. Consequently, from condition II in Definition 2.4 of RMPFr's, a right matrix Padé table for normal power series may be made unique by insisting that the constant term, V(0), in the denominator of any Padé fraction be the identity matrix.

Given the power series (2.4) and any non-negative integers m and n, we introduce a sequence of points

$$(m_0,n_0),\ (m_1,n_1),\ (m_2,n_2),\ \cdots \tag{3.1}$$

in the right matrix Padé table satisfying

$$(m_0,n_0) = \begin{cases} (m-n-1,-1) & , m \geq n \ , \\ (-1,n-m-1) & , m < n \ , \end{cases} \tag{3.2}$$

and

$$(m_{i+1},n_{i+1}) = (m_i + s_i\ ,n_i + s_i\ ),\ i = 0, 1, 2, \cdots , \tag{3.3}$$

where $1 \leq s_i \leq \infty$. Observe that

$$m_i - n_i = m - n \ , \ i = 0, 1, 2, \cdots , \tag{3.4}$$

and consequently the sequence (3.1) lies along the m-n off-diagonal path of the right matrix Padé table. In (3.3), the s_i are selected so that

$$d_{m_i,n_i} \neq 0 \tag{3.5}$$

and

$$d_{(m_i+j),(n_i+j)} = 0 \ , \tag{3.6}$$

for $j = 1, 2, \cdots , s_i - 1$.

For i = 1, 2, ..., let $(U_i(z), V_i(z), W_i(z))$ be the unique RMPFr (c.f., Theorem 2.5) of type (m_i, n_i) for (A(z),B(z)). Then $\left[V_i \ , U_i \ \right]^t$ satisfies

$$S_{m_i,n_i} \cdot \begin{bmatrix} V_i \\ U_i \end{bmatrix} = 0. \tag{3.7}$$

and

$$A(z) \cdot V_i(z) + B(z) \cdot U_i(z) = z^{m_i + n_i + 1} W_i(z). \tag{3.8}$$

Generalizing the notions of Cabay and Kossowski [17] , we introduce

Definition 3.1. The sequence

$$\left\{ W_i(z) \right\}, i = 1, 2, \cdots , \tag{3.9}$$

is called the **Power Series Remainder Sequence** for the pair (A(z),B(z)). The sequence of pairs

$$\left\{ (U_i(z), V_i(z)) \right\}, i = 1, 2, \cdots , \tag{3.10}$$

is called the corresponding **cofactor sequence.** The integer pairs $\{(m_i , n_i)\}$ are called **nonsingular nodes** along the m - n off-diagonal path of the Padé table for (A(z),B(z)).

We note that each term of a power series remainder sequence is unique up to multiplication on the right by a nonsingular matrix. This is also true for each term of the corresponding cofactor sequence.

Initially, when $m \geq n$, observe that $m_1 = m - n$ and $n_1 = 0$ (i.e., $s_0 = 1$), because in (2.22) the non-singularity of b_0 implies that $d_{(m-n),0} \neq 0$. Thus, $V_1(z)$ is some arbitrary nonsingular matrix from K and, using (3.6), $U_1(z)$ can be obtained by solving

$$\begin{bmatrix} b_0 & & \\ \cdot & \cdot & \\ \cdot & & \cdot \\ b_{m_1} & \cdots & b_0 \end{bmatrix} U_1 = - \begin{bmatrix} a_0 \\ \cdot \\ \cdot \\ a_{m_1} \end{bmatrix} V_1 . \tag{3.11}$$

That is, $U_1(z)$ can be obtained by multiplying the first m_1+1 terms of the quotient power series $B^{-1}(z) \cdot A(z)$ on the right by $-V_1(z)$.

Initially, when $m < n$, depending on a_0 there are two cases to consider. The simple case, when $det(a_0) \neq 0$, yields

$$d_{0,(n-m)} = \det \begin{bmatrix} a_0 & & \\ & \cdot & \\ & & \cdot \\ a_{n-m-1} & \cdots & a_0 \end{bmatrix} \neq 0 . \tag{3.12}$$

Thus, $s_0 = 1$, $m_1 = 0$, and $n_1 = n - m$. Then, the RMPFr $(U_1(z), V_1(z), W_1(z))$ of type (m_1, n_1) is determined by setting $U_1(z)$ to be an arbitrary nonsingular matrix from K and then solving

$$\begin{bmatrix} a_0 & & \\ \cdot & \cdot & \\ \cdot & & \cdot \\ a_{n_1} & \cdots & a_0 \end{bmatrix} V_1 = - \begin{bmatrix} b_0 \\ \cdot \\ \cdot \\ b_{n_1} \end{bmatrix} U_1 . \tag{3.13}$$

That is, when $m < n$ and $det(a_0) \neq 0$, $V_1(z)$ can be obtained by multiplying the first n_1+1 terms of the quotient power series $A^{-1}(z) \cdot B(z)$ on the right by $-U_1(z)$.

When $m < n$ and $det(a_0) = 0$, we must first determine the smallest positive integer s_0 (i.e., the smallest $m_1 = m_0+s_0$ and $n_1 = n_0+s_0$) so that $d_{m_1,n_1} \neq 0$. Then, $(U_1(z), V_1(z))$ is obtained by solving

$$S_{m_1,n_1} \cdot \begin{bmatrix} V_1 \\ U_1 \end{bmatrix} = 0 . \tag{3.14}$$

In the next section, we give an algorithm which computes a RMPFo of type (m,n) for (A(z),B(z)) by performing a sequence of the above types of initializations (albeit, each for different power series).

When the power series pair (A(z),B(z)) is normal, only the initializations corresponding to (3.11) and (3.13) are required. Thus, for normal power series $s_i = 1$ for all i, and the algorithm reduces to a sequence of truncated power series divisions.

There are also some non-normal power series that share this property. For each pair of integers m and n, let $r_{m,n}$ be the rank of the matrix $T_{m,n}$. Then normality is equivalent to

$$r_{m,n} = (m+n) \cdot p \tag{3.15}$$

for all m and n. A matrix power series pair (A(z), B(z)) is said to be **nearly-normal** (c.f., Labahn [32]) if, for all integers m and n,

$$r_{m,n} = k_{m,n} \cdot p \tag{3.16}$$

for some integer $k_{m,n}$. Clearly, every normal power series is also a nearly-normal power series. In addition, all scalar power series are nearly-normal.

For a nearly-normal power series pair (A(z),B(z)) it is easy to see that when a_0 is singular, then $a_0 = 0$. This follows from the observation that the rank of a_0 is just $r_{0,1}$, which, if it is not p, must be zero. Also, if $a_0 = \cdots = a_{k-1} = 0$ and $a_k \neq 0$, then a_k must be a nonsingular matrix for similar reasons. When $k > m$ this implies that there are no nonsingular nodes along the m - n off-diagonal path before and includ-

ing the node (m,n). Otherwise, when $k \le m$, the initialization (3.14) becomes

$$\begin{bmatrix} 0 & & & b_0 & & \\ & & & & \ddots & \\ a_{m_1} & & & \vdots & & b_0 \\ \vdots & & & \vdots & & \\ a_{m_1+s_1} & \cdots & a_{m_1} & b_{m_1+n_1} & \cdots & b_{n_1} \end{bmatrix} \cdot \begin{bmatrix} V_1 \\ \\ U_1 \end{bmatrix} = 0, \tag{3.17}$$

where $s_0 = k+1$, $m_1 = k$ and $n_1 = n-m+k$. Consequently the RMPFr $(U_1(z),V_1(z),W_1(z))$ of type (m_1,n_1) is obtained from (3.17) first by setting $U_1(z) = z^{m_1} \cdot U$, where U is any nonsingular matrix from K. Then, $V_1(z)$ is obtained by multiplying the first n_1+1 terms of the quotient power series $(z^{-m_1} \cdot A(z))^{-1} \cdot B(z)$ on the right by -U. Thus, also for nearly-normal power series (and therefore also for all scalar power series), all initializations reduce to truncated power series divisions.

Corresponding to the Power Series Remainder Sequence, we introduce

Definition 3.2. The sequence

$$\left\{ (P_i(z),Q_i(z)) \right\}, \quad i = 1, 2, \cdots, \tag{3.18}$$

where $(P_i(z),Q_i(z),R_i(z))$ is a Padé form of type (m_i-1, n_i-1) for (A(z),B(z)), is called a **predecessor sequence** of the power series remainder sequence.

The triple $(P_i(z),Q_i(z), R_i(z))$ satisfies the equation

$$A(z) \cdot Q_i(z) + B(z) \cdot P_i(z) = z^{m_i + n_i - 1} \cdot R_i(z) \tag{3.19}$$

by (2.6). Initially, when $s_0 = 1$, the predecessor pair $(P_1(z),Q_1(z))$ is defined to be

$$P_1(z) = z^{m-n-1} I_p \ , \text{ and } \ Q_1(z) = 0 \tag{3.20}$$

if $m \ge n$, and

$$P_1(z)=0\ ,\ Q_1(z)=z^{n-m-1}I_p \tag{3.21}$$

if $m < n$. Both (3.20) and (3.21) satisfy (3.19), and therefore condition II for RMPFo's, but not necessarily conditions I and III.

Theorem 3.3: For i = 1, 2, ... , the predecessors $(P_i(z),Q_i(z))$ are unique up to right multiplication by a nonsingular matrix from K. In addition, the leading term of the residual, $R_i(0)$, is nonsingular.

Proof: The results are an immediate consequence of (3.20), (3.21) and Theorem 2.6.

The main result of this section is

Theorem 3.4. For any positive integer k, (k - 1, k) is a nonsingular node in the Padé table for $(W_i(z),R_i(z))$ if and only if $(m_i+k\ ,\ n_i+k)$ is a nonsingular node in the Padé table for (A(z), B(z)).

Proof: Let M_{11}, M_{21}, M_{12}, and M_{22} be matrices of dimension $p(n_i+k)$ x pk, $p(m_i+k)$ x pk, $p(n_i+k)$ x p(k-1), and $p(m_i+k)$ x p(k-1), respectively, defined by

$$M_{11}=\begin{bmatrix} v_0 & & \\ v_{n_i} & \ddots & \\ & v_0 & \\ & \ddots & \\ & & v_{n_i} \end{bmatrix}\ ,\ M_{12}=\begin{bmatrix} 0 & & \\ 0 & & \\ q_0 & & \\ \vdots & \ddots & \\ q_{n_i-1} & q_0 & \\ & \ddots & \\ & q_{n_i-1} & \end{bmatrix}\ , \tag{3.22}$$

$$M_{21}=\begin{bmatrix} u_0 & & \\ u_{m_i} & \ddots & \\ & u_0 & \\ & \ddots & \\ & & u_{m_i} \end{bmatrix}\ ,\ M_{22}=\begin{bmatrix} 0 & & \\ 0 & & \\ p_0 & & \\ \vdots & \ddots & \\ p_{m_i-1} & p_0 & \\ & \ddots & \\ & p_{m_i-1} & \end{bmatrix}\ . \tag{3.23}$$

Also, let M be

$$M = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix}. \tag{3.24}$$

If we set

$$R_i(z) = \sum_{j=0}^{\infty} r_j z^j \ , \ with \ \ det(r_0) \neq 0, \ \text{ and } \ W_i(z) = \sum_{j=0}^{\infty} w_j z^j \ , \tag{3.25}$$

then, from (3.8) and (3.19), it follows that

$$T_{(m_i+k),(n_i+k)} \cdot M = \left[\begin{array}{ccc|ccc} 0 & \cdots & 0 & 0 & \cdots & 0 \\ & & & & & \\ & & & & & \\ 0 & \cdots & 0 & 0 & \cdots & 0 \\ w_0 & & & r_0 & & \\ \cdot & \cdot & & \cdot & & \\ \cdot & & w_0 & \cdot & & r_0 \\ \cdot & & \cdot & \cdot & & \cdot \\ w_{2k-2} & \cdots & w_{k-1} & r_{2k-2} & \cdots & r_k \end{array}\right]$$

$$= \begin{bmatrix} 0 \\ T'_{(k-1),k} \end{bmatrix}, \tag{3.26}$$

where **0** represents a zero matrix of size $p(m_i+n_i+1)$ x $p(2k-1)$ and

$$T'_{(k-1),k} = \left[\begin{array}{ccc|ccc} w_0 & & & r_0 & & \\ & \cdot & & & \cdot & \\ & & w_0 & & & r_0 \\ & & & & & \\ w_{2k-2} & \cdots & w_{k-1} & r_{2k-2} & \cdots & r_k \end{array}\right]. \tag{3.27}$$

We are now in a position to prove the theorem. Assume $T_{(m_i+k),(n_i+k)}$ is nonsingular. We show that $T'_{(k-1),k}$ is then also nonsingular. Let

$$X = \left[X_1 , \cdots , X_{2k-1} \right]' \tag{3.28}$$

be a p(2k-1) x 1 vector that satisfies

$$T'_{(k-1),k} \cdot X = 0 . \tag{3.29}$$

Since $T_{(m+k),(n+k)}$ is nonsingular, (3.26) implies

$$M \cdot X = 0. \tag{3.30}$$

From (3.30), we then obtain that

$$v_0 \cdot X_1 = 0 , \tag{3.31}$$

and consequently $X_1 = 0$, because v_0 is nonsingular. The first block equation from (3.29) then implies

$$r_0 \cdot X_{k+1} = 0 . \tag{3.32}$$

Thus, $X_{k+1} = 0$ because r_0 is nonsingular. In a similar fashion, it follows that $X_2 = 0$ and $X_{k+2} = 0$. Continuing in this way, we obtain that X = 0, that is, $T'_{(k-1),k}$ is nonsingular.

Conversely, suppose that $T'_{(k-1),k}$ is nonsingular. Let $X = (X_1 , \cdots , X_{2k-1})$ be a 1 x p(2k-1) vector, $Y = (Y_1 , \cdots , Y_{m+n})$ a 1 x p(m_i+n_i) vector and Z a 1 x p vector. Consider

$$(Z, Y, X) \cdot T_{(m+k),(n+k)} = 0. \tag{3.33}$$

Multiplying both sides of (3.33) on the right by M, and using equation (3.26), it follows that

$$X \cdot T'_{(k-1),k} = 0. \tag{3.34}$$

Since $T'_{(k-1),k}$ is nonsingular, then X = 0. Then, in (3.33), using block columns 2 through n_i+1 and block columns n_i+k+2 through to m_i+n_i+k+1 of $T_{(m+k),(n+k)}$, we obtain

$$Y \cdot T_{m,n} = 0 . \tag{3.35}$$

Since $T_{m,n}$ is nonsingular, Y = 0. Finally, block column n_i+k+1 of $T_{(m+k),(n+k)}$ now yields

$$Z \cdot b_0 = 0 . \tag{3.36}$$

Since b_0 is nonsingular, Z = 0. Hence, $T_{(m+k),(n+k)}$ is nonsingular.

Theorem 3.4 allows us to calculate nonsingular nodes of a pair of power series by calculating nonsingular nodes of the residual pair of power series. This gives us an iterative method of calculating nonsingular nodes.

Theorem 3.5: The cofactor and predecessor sequences for (A(z),B(z)) satisfies

$$\begin{bmatrix} U_{i+1}(z) & P_{i+1}(z) \\ V_{i+1}(z) & Q_{i+1}(z) \end{bmatrix} = \begin{bmatrix} U_i(z) & P_i(z) \\ V_i(z) & Q_i(z) \end{bmatrix} \cdot \begin{bmatrix} I & 0 \\ 0 & z^2 \cdot I \end{bmatrix} \cdot \begin{bmatrix} V'(z) & Q'(z) \\ U'(z) & P'(z) \end{bmatrix} \tag{3.37}$$

where $(U'(z),V'(z))$ are the numerator and denominator of the RMPFr of type (s_i-1, s_i) for $(W_i(z),R_i(z))$, and where $(P'(z),Q'(z))$ is its predecessor.

Proof: From Theorem 3.4, (s_i-1, s_i) is the first nonsingular node for $(W_i(z),R_i(z))$. Thus, $(U'(z),V'(z))$ satisfies

$$W_i(z)\cdot V'(z) + R_i(z)\cdot U'(z) = z^{2s}W'(z), \tag{3.38}$$

where $V'(0)$ is nonsingular and $W'(z)$ is some matrix power series which we show later is exactly $W_{i+1}(z)$. Then, using (3.8), (3.19), (3.37) and (3.38) we get

$$A(z)\cdot V_{i+1}(z) + B(z)\cdot U_{i+1}(z)$$

$$= A(z)\cdot\left\{V_i(z)\cdot V'(z) + z^2 Q_i(z)\cdot U'(z)\right\} + B(z)\cdot\left\{U_i(z)\cdot V'(z) + z^2 P_i(z)\cdot U'(z)\right\}$$

$$= \left\{A(z)\cdot V_i(z) + B(z)\cdot U_i(z)\right\}\cdot V'(z) + \left\{A(z)\cdot Q_i(z) + B(z)\cdot P_i(z)\right\}\cdot z^2\cdot U'(z)$$

$$= z^{m+n+1}\cdot\left\{W_i(z)\cdot V'(z) + R_i(z)\cdot U'(z)\right\}$$

$$= z^{(m_i+s_i)+(n_i+s_i)+1} \cdot W'(z) . \tag{3.39}$$

Thus, $W'(z) = W_{i+1}(z)$ and condition II for a RMPFo is satisfied.

To verify condition I, expanding (3.37) gives

$$U_{i+1}(z) = U_i(z) \cdot V'(z) + z^2 P_i(z) \cdot U'(z) , \tag{3.40}$$

so that

$$\begin{aligned} \partial(U_{i+1}(z)) &\le max(m_i + s_i , 2 + (m_i - 1) + s_i - 1) \\ &= m_i + s_i . \end{aligned} \tag{3.41}$$

Similarly,

$$\partial(V_{i+1}(z)) \le n_i + s_i . \tag{3.42}$$

Finally, to verify condition II for RMPFr, observe that

$$V_{i+1}(0) = V_i(0) \cdot V'(0) , \tag{3.43}$$

and, consequently, $V_{i+1}(0)$ is invertible since both $V_i(0)$ and $V'(0)$ are invertible.

Thus, $(U_{i+1}(z), V_{i+1}(z), W'(z))$ is a RMPFr of type (m_i+s_i, n_i+s_i) for (A(z),B(z)).

A similar argument shows that $(P_{i+1}(z), Q_{i+1}(z), R'(z))$ is the RMPFo of type $(m_{i+1}-1, n_{i+1}-1)$ and hence we have our predecessor.

3.2. The Algorithm

Given non-negative integers m and n, the algorithm MPADE below makes use of Theorem 3.5 to compute the cofactor and predecessor sequences (3.10) and (3.18), respectively. Thus, intermediate results available from MPADE include those RMPFr's $(U_i(z), V_i(z), W_i(z))$ for (A(z),B(z)) at all the nonsingular nodes (m_i, n_i), i=1,2, ... , k-1, smaller than (m,n), along the off-diagonal path $m_i - n_i$ = m-n. The output

gives results associated with the final node (m_k,n_k). If (m,n) is also a nonsingular node, then the output $(U_k(z),V_k(z),W_k(z))$ is a RMPFr of type (m,n) for (A(z),B(z)), and $(P_k(z),Q_k(z),R_k(z))$ is a RMPFo of type (m-1,n-1). If (m,n) is a singular node, then the output $(U_k(z),V_k(z),W_k(z))$ is simply a RMPFo of type (m,n) for (A(z),B(z)), and now $(P_k(z),Q_k(z),R_k(z))$ is set to be the RMPFr of type (m_{k-1}, n_{k-1}).

Note that, when (m,n) is not a nonsingular node, a simple modification of MPADE allows the computation of all RMPFo's of type (m,n) for (A(z),B(z)). It is only necessary to arrange to compute q columns of $[V'_k , U'_k]$, rather than p, in order to form a basis for the solution space of the equation in step 3.1 of MPADE. Here q (with $q \geq p$) is the number of linearly independent solutions of the system (2.11). From this basis, it is then possible to construct a p x p matrix V(z), and a corresponding U(z) and W(z), for which (U(z),V(z),W(z)) is a RMPFo of type (m,n) for (A(z),B(z)) and has the property that V(z) is an invertible matrix, assuming such a RMPFo exists. This enhancement is not included in MPADE primarily to simplify the presentation of the algorithm.

ALGORITHM (MPADE):

INPUT: A(z), B(z), m, n, and p, where

1) m and n are nonnegative integers, p is a positive integer,

2) A(z) and B(z) are p x p matrix power series with $det\,(B\,(0)) \neq 0$. Note that A(z) mod z^{m+n+1} and B(z) mod z^{m+n+1}, only, are required.

OUTPUT: Matrix polynomials $P_k(z)$, $Q_k(z)$, $U_k(z)$, $V_k(z)$, matrix power series $R_k(z)$, $W_k(z)$ and a boolean SUCCESS defined as follows:

1) If SUCCESS is true, then $(U_k(z),V_k(z),W_k(z))$ is a RMPFr of type (m,n) for (A(z),B(z)) and $(P_k(z),Q_k(z),R_k(z))$ is its predecessor;

2) If SUCCESS is false, then $(U_k(z),V_k(z),W_k(z))$ is a RMPFo of type (m,n) and $(P_k(z),Q_k(z),R_k(z))$ is a RMPFr at the largest nonsingular node along the m-n off-diagonal path.

Step 1: # Initialization #

If $m \geq n$

then set

1.1) $i \leftarrow 1$

1.2) $s_0 \leftarrow m - n$

1.3) $$\begin{bmatrix} m_1 \\ n_1 \end{bmatrix} = \begin{bmatrix} s_0 \\ 0 \end{bmatrix}$$

1.4) $$\begin{bmatrix} U_1(z) & P_1(z) \\ V_1(z) & Q_1(z) \end{bmatrix} = \begin{bmatrix} -B(z)^{-1} \cdot A(z) \bmod z^{s_0+1} & z^{s_0-1} \cdot I \\ I & 0 \end{bmatrix}$$

else set

1.5) $i \leftarrow 0$

1.6) $$\begin{bmatrix} m_0 \\ n_0 \end{bmatrix} = \begin{bmatrix} m-n \\ 0 \end{bmatrix}$$

1.7) $$\begin{bmatrix} U_0(z) & P_0(z) \\ V_0(z) & Q_0(z) \end{bmatrix} = \begin{bmatrix} 0 & z^{m-n-1} \cdot I \\ I & 0 \end{bmatrix}$$

Step 2: # Search for next nonsingular node #

2.1) $s_i \leftarrow 0$

2.2) $d \leftarrow 0$

2.3) Do while $n_i + s_i < n$ and $d = 0$

2.4) Set $s_i \leftarrow s_i + 1$

2.5) Compute the residual $W_i(z)$ such that

$$(A(z) \cdot V_i(z) + B(z) \cdot U_i(z)) \bmod z^{m_i+n_i+2s_i+1} = z^{m_i+n_i+1} \cdot W_i(z)$$

2.6) Compute the residual $R_i(z)$ such that

$$(A(z)\cdot Q_i(z) + B(z)\cdot P_i(z)) \bmod z^{m_i+n_i+2s_i-1} = z^{m_i+n_i-1}\cdot R_i(z)$$

2.7) Compute

$$d = det(T'_{(s_i-1),s_i}).$$

determined from the power series $W_i(z)$ and $R_i(z)$ (c.f., (3.27)).

2.8) End do

Step 3: # Compute RMPFr for residuals #

3.1) Solve

$$S'_{(s_i-1),s_i}\cdot\begin{bmatrix}V'\\U'\end{bmatrix} = 0,$$

where S' is the Sylvester matrix determined from $W_i(z)$ and $R_i(z)$

Step 4: # Compute predecessor for residuals #

4.1) If $s_i > 1$ and $d \neq 0$,

then solve

$$S'_{(s_i-2),(s_i-1)}\cdot\begin{bmatrix}Q'\\P'\end{bmatrix} = 0,$$

where again S' is the Sylvester matrix determined from $W_i(z)$ and $R_i(z)$

else set

4.2) $$\begin{bmatrix}Q'(z)\\P'(z)\end{bmatrix} = \begin{bmatrix}I\\0\end{bmatrix}$$

Step 5: # Advance along off-diagonal for Padé fractions #

5.1) $m_{i+1} \leftarrow m_i + s_i$

5.2) $n_{i+1} \leftarrow n_i + s_i$

5.3) $$\begin{bmatrix}U_{i+1}(z) & P_{i+1}(z)\\V_{i+1}(z) & Q_{i+1}(z)\end{bmatrix} \leftarrow \begin{bmatrix}U_i(z) & P_i(z)\\V_i(z) & Q_i(z)\end{bmatrix}\cdot\begin{bmatrix}I & 0\\0 & z^2 I\end{bmatrix}\cdot\begin{bmatrix}V'(z) & Q'(z)\\U'(z) & P'(z)\end{bmatrix}$$

5.4) $i \leftarrow i+1$

Step 6: # termination test #

If $n_i < n$

then go to step 2

Else

If $d \neq 0$

then SUCCESS ← true,

else SUCCESS ← false;

Exit.

Theorem 3.6: The MPADE algorithm is valid.

Proof: The argument is by induction on i.

When $m \geq n$, the initial step, i=1, to compute $(U_1(z),V_1(z))$ and $(P_1(z),Q_1(z))$ in step 1 is clearly valid. When $m < n$, the initialization in step 1 uses negative powers of z (invoked in order to simplify the algorithm presentation) and, consequently, $(U_0(z),V_0(z))$ and $(P_0(z),Q_0(z))$ do not satisfy the requirements of a RMPFo in a strict sense. However, it is a simple, but tedious, matter to show that one pass through MPADE yields $(U_1(z),V_1(z))$ and $(P_1(z),Q_1(z))$ which meet all the requirements for the first term in the cofactor and predecessor sequences (3.10) and (3.18), respectively (see the Appendix for details).

The inductive step along the nonsingular nodes of the off-diagonal path $m_i - n_i$ =m-n follows directly from Theorem 3.5 (compare (3.37) with step 5.3 of MPADE). Thus, when (m,n)=(m_k,n_k) is a nonsingular node, the proof is complete. On the other hand, when (m,n)=(m_k,n_k) is not a nonsingular node, arguments similar to those of Theorem 3.5 can be used to show that $(U'(z),V'(z),W'(z))$ is a RMPFo of type $(s_{k-1}-1,s_{k-1})$ for $(W_{k-1}(z),R_{k-1}(z))$ if and only if $(U_k(z),V_k(z),W_k(z))$ determined according to (3.37) is a

RMPFo of type (m_k, n_k) for (A(z),B(z)).

Remark 1. In the scalar case the MPADE algorithm reduces to the $O(n^2)$ off-diagonal algorithm of Cabay and Choi [16] . By reversing coefficients and with B(z)=-I, the anti-diagonal algorithm of McEliece and Shearer [35] is obtained.

Remark 2. For normal matrix power series, intermediate results are identical to those of Rissanen [41] . The algorithm of Rissanen does generalize to non-normal power series, but only in the scalar case (c.f., Rissanen [39]). It is not clear how to express this in the matrix case because of the problem of matrix singularity.

3.3. Complexity of the MPADE Algorithm

In assessing the costs of MPADE, it is assumed that classical algorithms are used for the multiplication of polynomials. Only the more costly steps are considered. For these steps, Table 5.1 below provides crude upper bounds on the number of multiplications in K performed during the i-th pass through MPADE.

Step	Bound on Number of Multiplications
2.5	$2p^3(m_i+n_i+2)(s_i+1)$
2.6	$2p^3(m_i+n_i+2)(s_i+1)$
2.7	$4p^3(s_i-1)^3$
3.1	$6p^3s_i^2$
4.1	$6p^3s_i^2$
5.3	$4p^3(m_i+n_i+2)(s_i+1)$

Table 3.6
Bounds on Operations per Step

In step 2.7 of MPADE, it is assumed that the Gaussian elimination method is used to obtain the LU decomposition of $T_{(s_i-1),s_i}$. In addition, it is assumed that Gaussian elimination is accompanied with border-

ing techniques. Thus, as s_i increases by 1 in step 2.4, the results of the previous pass through the **while** loop are used to achieve the current LU decomposition. The bound for step 2.7 in Table 3.6 assumes we do not take any advantage of the special nature of $T_{(s_i-1),s_i}$.

For step 3.1, it is assumed that the LU decomposition of $T_{(s_i-1),s_i}$ from step 2.7, is used to simplify the triangularization of $S_{(s_i-1),s_i}$. The solution $[V', U']$ is obtained finally by solving this triangularized $S_{(s_i-1),s_i}$. Similar observations apply to step 4.1.

An upper bound for the number of multiplications in K required by MPADE is obtained by summing the costs in Table 3.6 for i=0,1, ... ,k. We use the fact that

$$\sum_{i=0}^{k} s_i = m \text{ , if } m \geq n \text{ , and } \sum_{i=0}^{k} s_i = n \text{ , if } m < n \text{ .} \tag{3.44}$$

In addition,

$$\sum_{i=0}^{k} {m_i}^{\alpha} {s_i}^{\beta} \leq m^{\alpha+\beta} \text{ and } \sum_{i=0}^{k} {n_i}^{\alpha} {s_i}^{\beta} \leq n^{\alpha+\beta} . \tag{3.45}$$

Then, step 2.7 has a complexity of $O(p^3(m+n)^3)$ and the remaining steps a complexity of $O(p^3(m+n)^2)$, at worst.

Theorem 3.7 When the quotient power series $B(z)^{-1}A(z)$ is nearly-normal, the MPADE algorithm calculates a RMPFr of type (m,n) with a complexity of $O(p^3(m+n)^2)$.

Proof: When (A(z),B(z)) is nearly-normal s_i is often larger than one, but the matrix $T_{(s_i-1),s_i}$ is always in triangular form (c.f., Labahn [32]). Thus the cost of step 2.7 of MPADE is zero. Using table 3.6 along with (3.44) and (3.45) gives the result.

Corollary 3.8 When the quotient matrix power series $B(z)^{-1}A(z)$ is normal, the MPADE algorithm calculates a RMPFr of type (m,n) with complexity $O(p^3(m+n)^2)$.

Proof: Corollary 3.8 follows from Theorem 3.7 since all normal matrix power series are nearly-normal.

Corollary 3.9 For scalar quotient power series $B(z)^{-1}A(z)$, the MPADE algorithm calculates a RMPFr of type (m,n) with a complexity of $O((m+n)^2)$.

Proof: Corollary 3.9 follows from Theorem 3.7 since all scalar power series are nearly-normal.

When the matrix power series is not nearly-normal or normal the complexity of MPADE depends on the distribution of the nonsingular nodes along the particular off-diagonal. However, even one nonsingular node can reduce the cost substantially. For example, if the midpoint (n/2,n/2) is the only nonsingular node along the main diagonal then MPADE requires $1/3(n/2)^3 + 1/3(n/2)^3 = 1/12n^3$ operations, versus $1/3n^3$ operations needed for Gaussian elimination, a cost saving of a factor of 4. Algorithms that require normality, on the other hand, break down when even one node is singular.

Chapter 4

Inverse Formulae of Block Hankel Matrices

4.1. Introduction

Let

$$H_{m,n} = \begin{bmatrix} a_{m-n+1} & \cdots & a_m \\ \vdots & & \vdots \\ a_m & \cdots & a_{m+n-1} \end{bmatrix} \tag{4.1}$$

be a nonsingular block Hankel matrix with coefficients from the ring of p x p matrices over a field (with $a_i = 0$ for negative i). [1] The primary contribution of this chapter is a set of new closed formulae for $H_{m,n}^{-1}$. We require only that $H_{m,n}$ be nonsingular. When p=1, one of the formulae agrees with that obtained by Choi [18].

The representations for $H_{m,n}^{-1}$ depend on the concept of a matrix Padé form developed in chapter 2 for the matrix polynomial

$$A(z) = \sum_{i=0}^{m+n} a_i z^i. \tag{4.2}$$

Central to our approach are commutativity relationships which are shown to exist between certain matrix Padé forms. These commutativity relationships allow us to overcome the limitations imposed when using bordering techniques (which require additional nonsingularity conditions) that are the traditional methods for obtaining closed inverse formulae. Indeed, the conditions that we impose are both necessary and sufficient for the existence of an inverse.

A major advantage of a closed inverse formula is that it allows for efficient algorithms to calculate the inverses of Hankel matrices. This efficiency comes both in the cost complexity of calculating the

[1] All results hold, with minor modifications, for block Toeplitz matrices.

inverse and also in the amount of storage required for the final result.

When discussing Padé forms in the context of Hankel matrices, we use a minor change of notation from the previous chapters. A subscript on the numerator and denominator of a matrix Padé will now signify the type (i.e., maximal allowed degree) of the Padé form. This is done for purposes of simplification and clarification of our presentation.

4.2. Nonsingular Block Hankel Matrices and Matrix Padé Forms

For any block Hankel matrix (4.1), there is an associated matrix polynomial A(z) given by (4.2). The nonsingularity of a block Hankel matrix, is closely related to the existence of certain matrix Padé forms for the pair (A(z),-I). Notice that, for this special case, Condition II of Definition 2.1 becomes

$$\begin{bmatrix} & & a_0 \\ & \cdot\ \cdot & \\ \cdot & \cdot & \\ a_{m-n} & \cdots & a_m \end{bmatrix} \cdot \begin{bmatrix} v_n \\ \cdot \\ \cdot \\ v_0 \end{bmatrix} = \begin{bmatrix} u_0 \\ \cdot \\ \cdot \\ u_m \end{bmatrix} \tag{4.3}$$

and

$$\begin{bmatrix} a_{m-n+1} & \cdots & a_{m+1} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ a_m & \cdots & a_{m+n} \end{bmatrix} \cdot \begin{bmatrix} v_n \\ \cdot \\ \cdot \\ v_0 \end{bmatrix} = \begin{bmatrix} 0 \\ \cdot \\ \cdot \\ 0 \end{bmatrix}. \tag{4.4}$$

The matrix polynomial $V_n(z)$ can be determined by solving (4.4), that is solving a system of equations with the block Hankel matrix (4.1) as the coefficient matrix. Once $V_n(z)$ has been obtained, then $U_m(z)$ can be obtained from (4.3).

For purposes of presentation, we use the following convention. For a given pair of positive integers (m,n), let the triples $(U_m(z),V_n(z),W(z))$ and $(U_m^*(z),V_n^*(z),W^*(z))$ denote a RMPFo and a LMPFo, respectively, of type (m,n) for A(z). For the same (m,n), a RMPFo and a LMPFo of type (m-1,n-1) for A(z) will be represented, respectively, by $(P_{m-1}(z),Q_{n-1}(z),R(z))$ and $(P_{m-1}^*(z),Q_{n-1}^*(z),R^*(z))$. For these

Padé forms, collectively, condition II of Definition 2.1 becomes

$$A(z)V_n(z) - U_m(z) = z^{m+n+1} \cdot W(z), \tag{4.5}$$

$$V_n^*(z)A(z) - U_m^*(z) = z^{m+n+1} \cdot W^*(z), \tag{4.6}$$

$$A(z)Q_{n-1}(z) - P_{m-1}(z) = z^{m+n-1} \cdot R(z), \tag{4.7}$$

and

$$Q_{n-1}^*(z)A(z) - P_{m-1}^*(z) = z^{m+n-1} \cdot R^*(z). \tag{4.8}$$

In section 3, in the case that $H_{m,n}$ is nonsingular, the inverse is given in terms of these four matrix Padé forms.

Theorem 4.1: For a pair of positive integers (m,n), the following statements are equivalent:

$$det(H_{m,n}) \neq 0, \tag{4.9}$$

$$det(r_0) \neq 0 \quad \text{and} \quad det(v_0) \neq 0, \tag{4.10}$$

$$det(r_0^*) \neq 0 \quad \text{and} \quad det(v_0^*) \neq 0. \tag{4.11}$$

Proof: That (4.9) implies (4.10) and (4.9) implies (4.11) was proved in Theorems 2.5 and 2.6 since $det(T_{m,n}) \neq 0$ if and only if $det(H_{m,n})$, and so we show only the converses here. To see that (4.10) implies (4.9), let $X = (x_1, \cdots, x_n)$ be a row vector of length np and suppose that

$$X \cdot H_{m,n} = 0. \tag{4.12}$$

We shall show that $X = 0$. We accomplish this by showing that (4.12) implies that $x_n = 0$ and

$$(0, x_1, \cdots, x_{n-1}) \cdot H_{m,n} = 0. \tag{4.13}$$

By repeated application of this property, it then follows that $x_{n-1} = \cdots = x_1 = 0$, and so $X = 0$.

First observe that equating coefficients of z^i, for $m+1 \leq i \leq m+n$, in (4.5) yields

$$H_{m,n}\cdot\begin{bmatrix} v_n \\ . \\ . \\ v_1 \end{bmatrix} = -\begin{bmatrix} a_{m+1} \\ . \\ . \\ a_{m+n} \end{bmatrix} v_0 , \tag{4.14}$$

where v_0 is invertible since we are assuming statement (4.10). Similarly, equating coefficients of z^i, for $m \le i \le m+n-1$, in (4.7) yields

$$H_{m,n}\cdot\begin{bmatrix} q_{n-1} \\ . \\ . \\ q_0 \end{bmatrix} = \begin{bmatrix} 0 \\ . \\ 0 \\ r_0 \end{bmatrix}, \tag{4.15}$$

where, by assumption, r_0 is invertible. From (4.12) and (4.15), it follows that

$$x_n \cdot r_0 = X \cdot \begin{bmatrix} 0 \\ . \\ 0 \\ r_0 \end{bmatrix} = X \cdot H_{m,n} \cdot \begin{bmatrix} q_{n-1} \\ . \\ . \\ q_0 \end{bmatrix} = 0. \tag{4.16}$$

Since r_0 is invertible, it then follows that $x_n = 0$.

Having shown that $x_n = 0$, equation (4.12) then yields

$$(x_1, \cdots, x_{n-1}) \cdot \begin{bmatrix} a_{m-n+2} & \cdots & a_m \\ . & & . \\ . & & . \\ a_m & \cdots & a_{m+n-2} \end{bmatrix} = 0. \tag{4.17}$$

But, from (4.12) and (4.14), we have

$$(x_1, \cdots, x_{n-1}, 0) \cdot \begin{bmatrix} a_{m+1} \\ . \\ . \\ a_{m+n} \end{bmatrix} \cdot v_0 = X \cdot H_{m,n} \begin{bmatrix} v_n \\ . \\ . \\ v_1 \end{bmatrix} = 0. \tag{4.18}$$

Since v_0 is invertible, (4.18) implies that

$$(x_1, \cdots, x_{n-1}) \cdot \begin{bmatrix} a_{m+1} \\ . \\ . \\ a_{m+n-1} \end{bmatrix} = 0. \tag{4.19}$$

Equations (4.17) and (4.19) imply that

$$(x_1, \cdots, x_{n-1}) \begin{bmatrix} a_{m-n+2} & \cdots & a_{m+1} \\ \vdots & & \vdots \\ a_m & \cdots & a_{m+n-1} \end{bmatrix} = 0, \tag{4.20}$$

which is equivalent to (4.13).

Thus, we have shown that (4.10) implies (4.9). A similar argument shows that (4.11) implies (4.9).

Note that both conditions in (4.10) and in (4.11) of Theorem 4.1 are necessary. Consider

$$A(z) = z^3 + 0 \cdot z^4 + \cdots . \tag{4.21}$$

Then a RMPFo, $(P_1(z), Q_2(z), R(z))$, of type (1,2) is given by

$$P_1(z) = 0, \quad Q_2(z) = z, \quad R(z) = 1 + \cdots . \tag{4.22}$$

But $\det(H_{2,3}) = 0$, so that $\det(r_0) \neq 0$ alone does not imply (4.9). Also, consider

$$A(z) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} z^2 + \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} z^4 + \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} z^5 + \cdots . \tag{4.23}$$

Then one RMPFo, $(U_2(z), V_3(z), W(z))$, of type (2,3) is given by

$$U_2(z) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} z^2, \quad V_3(z) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} z^2 . \tag{4.24}$$

But, here again, $\det(H_{2,3}) = 0$. Thus, $\det(v_0) \neq 0$ alone does not imply (4.9).

Theorem 4.1 has important computational significance since the singularity of $H_{m,n}$ can be detected simply by recognizing a singular r_0 or a singular v_0. If both r_0 and v_0 are nonsingular, then the results of chapter 2 implies that the matrix Padé forms identified by (4.5), (4.6), (4.7) and (4.8) are unique, except for the specification of the nonsingular matrices v_0, v_0^*, r_0, and r_0^*. As a consequence it can be assumed

without loss-of generality that

$$v_0 = v_0^* = r_0 = r_0^* = I. \tag{4.25}$$

This nonrestrictive assumption simplifies the presentation of subsequent results.

A key relationship between matrix Padé forms which enables, in sections 3 and 4, the presentation of the inverse of $H_{m,n}$, is given by

Lemma 4.2. Let $\det(H_{m,n}) \neq 0$. Then the matrix Padé forms identified by (4.5), (4.6), (4.7) and (4.8) and normalized according to (4.25) satisfy

$$\begin{bmatrix} Q_{n-1}^*(z) & -P_{m-1}^*(z) \\ -V_n^*(z) & U_m^*(z) \end{bmatrix} \cdot \begin{bmatrix} U_m(z) & P_{m-1}(z) \\ V_n(z) & Q_{n-1}(z) \end{bmatrix} = z^{m+n-1} \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}, \tag{4.26}$$

$$\begin{bmatrix} U_m(z) & P_{m-1}(z) \\ V_n(z) & Q_{n-1}(z) \end{bmatrix} \cdot \begin{bmatrix} Q_{n-1}^*(z) & -P_{m-1}^*(z) \\ -V_n^*(z) & U_m^*(z) \end{bmatrix} = z^{m+n-1} \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}, \tag{4.27}$$

$$\begin{bmatrix} R^*(z) & -Q_{n-1}^*(z) \\ -z^2 W^*(z) & V_n^*(z) \end{bmatrix} \cdot \begin{bmatrix} V_n(z) & Q_{n-1}(z) \\ z^2 W(z) & R(z) \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \tag{4.28}$$

and

$$\begin{bmatrix} V_n(z) & Q_{n-1}(z) \\ z^2 W(z) & R(z) \end{bmatrix} \cdot \begin{bmatrix} R^*(z) & -Q_{n-1}^*(z) \\ -z^2 W^*(z) & V_n^*(z) \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}. \tag{4.29}$$

Proof: Multiplying (4.5) on the left by $Q_{n-1}^*(z)$ and (4.8) on the right by $V_n(z)$, and subtracting the first from the second, we obtain

$$Q_{n-1}^*(z) \cdot U_m(z) - P_{m-1}^*(z) \cdot V_n(z) = z^{m+n-1} \cdot (R^*(z) V_n(z) - z^2 Q_{n-1}^*(z) W(z))$$

$$= z^{m+n-1} r_0^* v_0$$

$$= z^{m+n-1} I. \tag{4.30}$$

In (4.30), we have used the normalizing condition (4.25) and the fact that the left hand side, and consequently the right hand side, is a matrix polynomial of degree at most m+n-1.

Multiplying (4.5) on the left by $V_n^*(z)$ and (4.6) on the right by $V_n(z)$, and subtracting the second from the first, we obtain

$$-V_n^*(z) \cdot U_m(z) + U_m^*(z) \cdot V_n(z) = z^{m+n+1} \cdot (V_n^*(z) W(z) - W^*(z) V_n(z))$$

$$= 0. \tag{4.31}$$

In (4.31), the last equality is true because the left hand side and consequently the right hand side, is a matrix polynomial of degree at most m+n.

In a similar fashion, (4.7), (4.8) and (4.25) yield

$$Q_{n-1}^*(z) P_{m-1}(z) - P_{m-1}^*(z) Q_{n-1}(z) = z^{m+n-1} \cdot (Q_{n-1}^*(z) R(z) - R^*(z) Q_{n-1}(z))$$

$$= 0; \tag{4.32}$$

whereas, (4.6), (4.7) and (4.25) give

$$-V_n^*(z) P_{m-1}(z) + U_m^*(z) Q_{n-1}(z) = z^{m+n-1} \cdot (V_n^*(z) R(z) - z^2 W^*(z) Q_{n-1})$$

$$= z^{m+n-1} I. \tag{4.33}$$

Equations (4.30), (4.31), (4.32) and (4.33) together comprise (4.26). Equation (4.27) follows directly from (4.26), since matrix inverses are two sided.

Equations (4.30) also gives

$$z^{m+n-1} \cdot (R^*(z) V_n(z) - z^2 Q_{n-1}^*(z) W(z)) = z^{m+n-1} \cdot I, \tag{4.34}$$

from which we obtain

$$R^*(z)V_n(z)-z^2Q^*_{n-1}(z)W(z)=I. \tag{4.35}$$

Similarly, from equation (4.31), we obtain

$$V^*_n(z)W(z)-W^*(z)V_n(z)=0. \tag{4.36}$$

From (4.32), we obtain

$$Q^*_{n-1}(z)R(z)-R^*(z)Q_{n-1}(z)=0, \tag{4.37}$$

and (4.33) gives

$$V^*_n(z)R(z)-z^2W^*(z)Q_{n-1}=I. \tag{4.38}$$

Equations (4.35), (4.36), (4.37) and (4.38) comprise (4.28). As before, (4.29) follows from (4.28), since matrix inverses are two sided.

4.3. The Off-diagonal Inverse Formulae

The main result of this chapter is

Theorem 4.3. Let $H_{m,n}$ be the block Hankel matrix (4.1). If there are RMPFo's and LMPFo's of type (m-1,n-1) and (m,n) for A(z) satisfying the normalizing condition (4.25), then $H_{m,n}$ is nonsingular with inverse

$$H^{-1}_{m,n}=\begin{bmatrix} v_{n-1} & & v_0 \\ & \iddots & \\ v_0 & & \end{bmatrix}\begin{bmatrix} q^*_{n-1} & & q^*_0 \\ & \iddots & \\ & & q^*_{n-1} \end{bmatrix}-\begin{bmatrix} q_{n-2} & & q_0\,0 \\ & \iddots & \\ q_0 & & \\ 0 & & \end{bmatrix}\begin{bmatrix} v^*_n & & v^*_1 \\ & \iddots & \\ & & v^*_n \end{bmatrix}, \tag{4.38}$$

or, equivalently,

$$H^{-1}_{m,n}=\begin{bmatrix} q_{n-1} & & \\ & \iddots & \\ q_0 & & q_{n-1} \end{bmatrix}\begin{bmatrix} v^*_{n-1} & & v^*_0 \\ & \iddots & \\ & & v^*_0 \end{bmatrix}-\begin{bmatrix} v_n & & \\ & \iddots & \\ v_1 & & v_n \end{bmatrix}\begin{bmatrix} q^*_{n-2} & & q^*_0\,0 \\ & \iddots & \\ & & q^*_0 \\ & & 0 \end{bmatrix}. \tag{4.39}$$

Proof: Using

$$U_m(z)Q_{n-1}^*(z) - P_{m-1}(z)V_n^*(z) = z^{m+n-1}I,$$

which is from (4.27), we can equate coefficients of z^i, $m \le i \le m+n-1$, to obtain

$$\begin{bmatrix} u_m & & u_{m-n+1} \\ & \ddots & \\ & & u_m \end{bmatrix} \cdot \begin{bmatrix} q_{n-1}^* & & q_0^* \\ & \ddots & \\ & & q_{n-1}^* \end{bmatrix} - \begin{bmatrix} p_{m-1} & & p_{m-n} \\ & \ddots & \\ & & p_{m-1} \end{bmatrix} \cdot \begin{bmatrix} v_n^* & & v_1^* \\ & \ddots & \\ & & v_n^* \end{bmatrix} = I. \tag{4.40}$$

Similarly,

$$V_n(z)Q_{n-1}^*(z) - Q_{n-1}(z)V_n^*(z) = 0,$$

also from (4.27), yields

$$\begin{bmatrix} & & v_n \\ & \ddots & \\ v_n & & v_1 \end{bmatrix} \begin{bmatrix} q_{n-1}^* & & q_0^* \\ & \ddots & \\ & & q_{n-1}^* \end{bmatrix} - \begin{bmatrix} & & q_{n-1} \\ & \ddots & \\ q_{n-1} & & q_0 \end{bmatrix} \begin{bmatrix} v_n^* & & v_1^* \\ & \ddots & \\ & & v_n^* \end{bmatrix} = 0. \tag{4.41}$$

Now, from (4.7), we obtain

$$H_{m,n} \cdot \begin{bmatrix} q_{n-2} & & q_0 0 \\ & \ddots & \\ q_0 & & \\ 0 & & \end{bmatrix} = \begin{bmatrix} p_{m-1} & & p_{m-n} \\ & \ddots & \\ & & p_{m-1} \end{bmatrix} - H_{m-n,n} \cdot \begin{bmatrix} & & q_{n-1} \\ & \ddots & \\ q_{n-1} & & q_0 \end{bmatrix}. \tag{4.42}$$

Observe that, for $1 \le i, j \le n$, the (i,j) component in (4.42) is obtained by equating coefficients of $z^{m+i-j-1}$ in (4.7). Similarly, (4.5) yields

$$H_{m,n} \cdot \begin{bmatrix} v_{n-1} & & v_0 \\ & \ddots & \\ v_0 & & \end{bmatrix} = \begin{bmatrix} u_m & & u_{m-n+1} \\ & \ddots & \\ & & u_m \end{bmatrix} - H_{m-n,n} \cdot \begin{bmatrix} & & v_n \\ & \ddots & \\ v_n & & v_1 \end{bmatrix}. \tag{4.43}$$

Combining (4.42) and (4.43) and using (4.40) and (4.41), it then follows that

$$\tilde{H}_{m,n}\cdot\left\{\begin{bmatrix} v_{n-1} & & v_0 \\ & \ddots & \\ v_0 & & \end{bmatrix}\cdot\begin{bmatrix} q_{n-1}^* & & q_0^* \\ & \ddots & \\ & & q_{n-1}^* \end{bmatrix}-\begin{bmatrix} q_{n-2} & \cdots & q_0\,0 \\ & \ddots & \\ q_0 & & \\ 0 & & \end{bmatrix}\cdot\begin{bmatrix} v_n^* & & v_1^* \\ & \ddots & \\ & & v_n^* \end{bmatrix}\right\}$$

$$=\left\{\begin{bmatrix} u_m & & u_{m-n+1} \\ & \ddots & \\ & & u_m \end{bmatrix}-H_{m-n,n}\cdot\begin{bmatrix} & & v_n \\ & \ddots & \\ v_n & & v_1 \end{bmatrix}\right\}\cdot\begin{bmatrix} q_{n-1}^* & & q_0^* \\ & \ddots & \\ & & q_{n-1}^* \end{bmatrix}$$

$$-\left\{\begin{bmatrix} p_{m-1} & & p_{m-n} \\ & \ddots & \\ & & p_{m-1} \end{bmatrix}-H_{m-n,n}\begin{bmatrix} & & q_{n-1} \\ & \ddots & \\ q_{n-1} & & q_0 \end{bmatrix}\right\}\cdot\begin{bmatrix} v_n^* & & v_1^* \\ & \ddots & \\ & & v_n^* \end{bmatrix}$$

$$=\left\{\begin{bmatrix} u_m & & u_{m-n+1} \\ & \ddots & \\ & & u_m \end{bmatrix}\cdot\begin{bmatrix} q_{n-1}^* & & q_0^* \\ & \ddots & \\ & & q_{n-1}^* \end{bmatrix}-\begin{bmatrix} p_{m-1} & & p_{m-n} \\ & \ddots & \\ & & p_{m-1} \end{bmatrix}\cdot\begin{bmatrix} v_n^* & & v_1^* \\ & \ddots & \\ & & v_n^* \end{bmatrix}\right\}$$

$$-H_{m-n,n}\left\{\begin{bmatrix} & & v_n \\ & \ddots & \\ v_n & & v_1 \end{bmatrix}\cdot\begin{bmatrix} q_{n-1}^* & & q_0^* \\ & \ddots & \\ & & q_{n-1}^* \end{bmatrix}-\begin{bmatrix} & & q_{n-1} \\ & \ddots & \\ q_{n-1} & & q_0 \end{bmatrix}\cdot\begin{bmatrix} v_n^* & & v_1^* \\ & \ddots & \\ & & v_n^* \end{bmatrix}\right\}$$

$$=I. \tag{4.44}$$

Thus, $H_{m,n}$ is nonsingular with inverse given by (4.38).

The second formula (4.39) for the inverse is proved using (4.6), (4.8) and the second column of (4.27).

Remark 1. In the scalar case, formula (4.38) was first obtained by Choi [18] .

Remark 2: The assumptions of Theorem 4.3 can be equivalently replaced by the requirement that we obtain solutions to

$$H_{m,n}\cdot\left[q_{n-1},\cdots,q_0\right]^t=\left[0,\cdots,0,I\right]^t, \tag{4.45}$$

$$\left[q_{n-1}^*,\cdots,q_0^*\right]\cdot H_{m,n}=\left[0,\cdots,0,I\right], \tag{4.46}$$

$$H_{m,n}\cdot\left[v_n,\cdots,v_1\right]^t=-\left[a_{m+1},\cdots,a_{m+n-1},a_{m+n}\right]^t, \tag{4.47}$$

and

$$\left[v_n^*,\cdots,v_1^*\right]\cdot H_{m,n}=-\left[a_{m+1},\cdots,a_{m+n-1},a_{m+n}\right], \tag{4.48}$$

where a_{m+n} can be any p x p matrix. Equations (4.47) and (4.48) are block versions of the Yule-Walker equations.

4.4. The Anti-diagonal Inverse Formulae

Theorem 4.3 provides inverse formulae for the block Hankel matrix $H_{m,n}$ in terms of RMPFo and LMPFo of type (m-1,n-1) and (m,n) for the associated matrix polynomial A(z). There are some algorithms (c.f., [8] , [35] , [44]) that calculate Padé forms along an anti-diagonal, rather than along an off-diagonal path of the Padé table. For this reason, it is useful to provide inverse formulae in terms of RMPFo's and LMPFo's of type (m-1,n) and (m,n-1) for A(z).

Let $(E_m(z),F_{n-1}(z),G(z))$ and $(E_m^*(z),F_{n-1}^*(z),G^*(z))$ be a RMPFo and a LMPFo, respectively, of type (m,n-1) for A(z). Also, let $(B_{m-1}(z),C_n(z),D(z))$ and $(B_{m-1}^*(z),C_n^*(z),D^*(z))$ be a RMPFo and a LMPFo, respectively, of type (m-1,n) for A(z). Then, the following equations are satisfied:

$$A(z)F_{n-1}(z) - E_m(z) = z^{m+n}G(z), \tag{4.49}$$

$$F_{n-1}^*(z)A(z) - E_m^*(z) = z^{m+n}G^*(z), \tag{4.50}$$

$$A(z)C_n(z) - B_{m-1}(z) = z^{m+n}D(z) \tag{4.51}$$

and

$$C_n^*(z)A(z) - B_{m-1}^*(z) = z^{m+n}D^*(z). \tag{4.52}$$

Corollary 4.1: Let $H_{m,n}$ be the block Hankel matrix (4.1). Then the following are equivalent:

$$det(H_{m,n}) \neq 0, \tag{4.53}$$

$$det(e_m) \neq 0 \text{ and } det(c_n) \neq 0, \tag{4.54}$$

$$det(e_m^*) \neq 0 \text{ and } det(c_n^*) \neq 0. \tag{4.55}$$

If any (and therefore all) of (4.53), (4.54) or (4.55) hold, then an inverse is given by

$$H_{m,n}^{-1} = \begin{bmatrix} c_n & & \\ & \ddots & \\ c_1 & & c_n \end{bmatrix}\begin{bmatrix} f_{n-1}^* & & f_0^* \\ & \ddots & \\ f_0^* & & \end{bmatrix} - \begin{bmatrix} 0 & & \\ f_{n-1} & & \\ & \ddots & \\ f_1 & f_{n-1} & 0 \end{bmatrix}\begin{bmatrix} c_{n-1}^* & & c_1^* \\ & \ddots & \\ c_0^* & & \end{bmatrix}, \tag{4.56}$$

or, equivalently,

$$H_{m,n}^{-1} = \begin{bmatrix} f_{n-1} & & f_0 \\ & \ddots & \\ f_0 & & \end{bmatrix}\begin{bmatrix} c_n^* & & c_1^* \\ & \ddots & \\ & & c_n^* \end{bmatrix} - \begin{bmatrix} c_{n-1} & & c_0 \\ & \ddots & \\ c_0 & & \end{bmatrix}\begin{bmatrix} 0 & f_{n-1}^* & f_1^* \\ & & \ddots \\ & & f_{n-1}^* \\ & & 0 \end{bmatrix}, \tag{4.57}$$

where we have normalized the Padé forms so that

$$e_m = e_m^* = c_n = c_n^* = I. \tag{4.58}$$

Proof: Let $a_i^\# = a_{2m-i}$, for $0 \le i = m+n$, and define a truncated power series $A^\#(z) = \sum_{i=0}^{m+n} a_i^\# z^i$.

Observe that, if

$$H_{m,n}^\# = \begin{bmatrix} a_{m-n+1}^\# & \cdots & a_m^\# \\ . & & . \\ a_m^\# & \cdots & a_{m+n-1}^\# \end{bmatrix}, \tag{4.59}$$

then

$$H_{m,n}^\# = J \cdot H_{m,n} \cdot J, \tag{4.60}$$

where

$$J = \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix}.$$

Equating coefficients of z^i, for $m \le i \le m+n-1$, in (4.49) gives

$$H_{m,n} \begin{bmatrix} f_{n-1} \\ . \\ f_0 \end{bmatrix} = \begin{bmatrix} e_m \\ 0 \\ . \\ 0 \end{bmatrix}. \tag{4.61}$$

From (4.60) and (4.61), it then follows that

$$H_{m,n}^\# \begin{bmatrix} f_0 \\ . \\ f_{n-1} \end{bmatrix} = \begin{bmatrix} 0 \\ . \\ 0 \\ e_m \end{bmatrix}. \tag{4.62}$$

Thus,

$$Q_{n-1}(z) = \sum_{i=0}^{n-1} f_{n-1+i} z^i \tag{4.63}$$

is a right denominator of type (m-1,n-1) for $A^{\#}(z)$. Similarly, (4.50) yields

$$\left[f_0^*, \cdots, f_{n-1}^*\right] H_{m,n}^{\#} = \left[0, \cdots, 0, e_m^*\right] \tag{4.64}$$

and so

$$Q_{n-1}^*(z) = \sum_{i=0}^{n-1} f_{n-1+i}^* z^i \tag{4.65}$$

is a left denominator of type (m-1,n-1) for $A^{\#}(z)$.

Next, from (4.51), we obtain

$$H_{m,n} \begin{bmatrix} c_{n-1} \\ . \\ c_0 \end{bmatrix} = \begin{bmatrix} a_{m-n} \\ . \\ a_{m-1} \end{bmatrix} c_n, \tag{4.66}$$

and so (4.60) then gives

$$H_{m,n}^{\#} \begin{bmatrix} c_0 \\ . \\ c_{n-1} \end{bmatrix} = \begin{bmatrix} a_{m+1}^{\#} \\ . \\ a_{m+n}^{\#} \end{bmatrix} c_n, \tag{4.67}$$

Thus,

$$V_n(z) = \sum_{i=0}^{n} c_{n-i} z^i \tag{4.68}$$

is a right denominator of type (m,n) for $A^{\#}(z)$. Similarly, (4.52) can be used to obtain

$$\left[c_0^*, \cdots, c_{n-1}^*\right] H_{m,n}^{\#} = c_n^* \left[a_{m+1}^{\#}, \cdots a_{m+n}^{\#}\right] \tag{4.69}$$

and so

$$V_n^*(z) = \sum_{i=0}^{n} c_{n-i}^* z^i \tag{4.70}$$

is a left denominator of type (m,n) for $A^{\#}(z)$. Since $\det(H_{m,n}^{\#}) \neq 0$ if and only if $\det(H_{m,n}) \neq 0$, the equivalence of (4.53), (4.54) and (4.55) now follows from the equivalence of (4.9), (4.10) and (4.11).

To prove (4.56), normalize according to (4.58) and substitute (4.63), (4.65), (4.68) and (4.70) into (4.38) to obtain

$$H_{m,n}^{\#^{-1}} = \begin{bmatrix} c_1 & & c_n \\ & \ddots & \\ c_n & & \end{bmatrix} \begin{bmatrix} f_0^* & & f_{n-1}^* \\ & \ddots & \\ & & f_0^* \end{bmatrix} - \begin{bmatrix} f_1 & & f_{n-1}\,0 \\ & \ddots & \\ f_{n-1} & & \\ 0 \ddots & & \end{bmatrix} \begin{bmatrix} c_0^* & & c_{n-1}^* \\ & \ddots & \\ & & c_0^* \end{bmatrix}, \tag{4.71}$$

By using (4.60), (4.56) follows immediately from (4.70). In a similar fashion, (4.57) can be obtained using (4.39).

Chapter 5

Other Inverse Formulae

5.1. Introduction

The special structure of Hankel matrices has resulted in a number of closed formulae for the inverse of $H_{m,n}$ other than those given in chapter 4. When p =1 (the scalar case) well known formulae of Gohberg and Semencul [23] gives $H_{m,n}^{-1}$ in terms of only the first and last columns of the inverse. Gohberg and Krupnik [24] give a formula for the inverse in terms of the last two columns of $H_{m,n}^{-1}$.

When p > 1, there are closed formulae due to Gohberg and Heinig [25] (these are actually given when the coefficients are from a noncommutative algebra). These are given provided the first and last columns together with the first and last rows of the inverse are known.

All of the above formulae depend on the ability to perform certain bordering operations that lend themselves well to matrices with a Hankel structure. However, these bordering operations require the imposition of certain additional restrictions on $H_{m,n}$. For the Gohberg-Krupnik formula, the matrix $H_{m-1,n-1}$ must also be nonsingular; whereas, for the Gohberg-Semencul and Gohberg-Heinig formulae, the matrix $H_{m,n-1}$ must be nonsingular.

In this chapter, we show that these formulae are simply special cases of the formulae from chapter 4. Indeed, when we add the condition that $H_{m,n-1}$ also be nonsingular, certain Frobenius-type identities for matrix Padé forms are used to show that our formulae yields the formulae of Gohberg and Heinig. On the other hand, when we add the condition that $H_{m-1,n-1}$ be nonsingular, a different set of Frobenius-type identities applied to our results yields inverse formulae, which in the scalar case corresponds to the Gohberg-Krupnik formulae. Finally, we show that additional formulae can be obtained using the results of chapter 4 along with various relationships existing in the Padé table of a power series.

5.2. The Inverse Formulae of Gohberg-Heinig

In this and the next section, we compare our off-diagonal inverse formulae (4.38) and (4.39) with other similar well-known formulae. In terms of matrix Padé forms of type (m-1,n-1) and (m,n-1), the inverse of $H_{m,n}$ is given by

Corollary 5.1. Let the matrix Padé forms identified by (4.7), (4.8), (4.49) and (4.50) be given. Then the following statements are equivalent:

$$det\,(H_{m,n-1}) \neq 0 \text{ and } det\,(H_{m,n}) \neq 0, \tag{5.1}$$

$$det\,(r_0) \neq 0 \text{ and } det\,(f_0) \neq 0, \tag{5.2}$$

$$det\,(r_0^*) \neq 0 \text{ and } det\,(f_0^*) \neq 0. \tag{5.3}$$

In addition, if any (and therefore all) of conditions (5.1), (5.2) or (5.3) are satisfied, then

$$H_{m,n}^{-1} = \begin{bmatrix} f_{n-1} & \cdots & f_0 \\ \vdots & & \\ f_0 & & \end{bmatrix} \begin{bmatrix} q_{n-1}^* & \cdots & q_0^* \\ & \ddots & \\ & & q_{n-1}^* \end{bmatrix} - \begin{bmatrix} q_{n-2} & q_0 & 0 \\ \vdots & \iddots & \\ q_0 & & \\ 0 & & \end{bmatrix} \begin{bmatrix} 0 & f_{n-1}^* & \cdots & f_1^* \\ & \ddots & \ddots & \\ & & & f_{n-1}^* \\ & & & 0 \end{bmatrix}, \tag{5.4}$$

where the Padé forms have been normalized by

$$r_0 = r_0^* = f_0 = f_0^* = I. \tag{5.5}$$

Proof: We first show that (5.1) implies (5.2). Since $\det(H_{m,n}) \neq 0$, Theorem 4.1 implies that $\det(r_0) \neq 0$. Since $\det(H_{m,n-1}) \neq 0$, Theorem 4.1 also implies that $\det(f_0) \neq 0$. Therefore (5.1) implies (5.2). In a similar fashion (5.1) implies (5.3).

To show that (5.2) implies (5.1), let

$$U_m(z) = E_m(z) - z \cdot P_{m-1}(z) r_0^{-1} \cdot g_0 \tag{5.6}$$

$$V_n(z) = F_{n-1}(z) - z \cdot Q_{n-1}(z) r_0^{-1} \cdot g_0 . \tag{5.7}$$

Then, $\partial(U_m(z)) \le m$ and $\partial(V_n(z)) \le n$. Also,

$$\begin{aligned} A(z)V_n(z) - U_m(z) &= \left\{A(z)F_{n-1}(z) - E_m(z)\right\} - z\left\{A(z)Q_{n-1}(z) - P_{m-1}(z)\right\}r_0^{-1}g_0 \\ &= z^{m+n}\left\{G(z) - R(z)r_0^{-1}g_0\right\} \\ &= z^{m+n+1}W(z), \end{aligned} \tag{5.8}$$

where

$$W(z) = z^{-1}\left\{G(z) - R(z)r_0^{-1}g_0\right\} \in D[[z]]. \tag{5.9}$$

Finally, the columns of $V_n(z)$ are linearly independent, since, from (5.7), $v_0 = f_0$, and by assumption f_0 is nonsingular. Thus, $(U_m(z), V_n(z), W(z))$ is a RMPFo of type (m,n) for A(z), satisfying $\det(v_0) \neq 0$. From Theorem 4.1, it follows that $H_{m,n}$ is nonsingular since both $\det(r_0) \neq 0$ and $\det(v_0) \neq 0$. To see that $H_{m,n-1}$ is also nonsingular, observe that

$$\begin{bmatrix} a_{m-n+1} & \cdots & a_m \\ \cdot & & \cdot \\ \cdot & & \cdot \\ a_m & \cdots & a_{m+n-1} \end{bmatrix} \begin{bmatrix} I & & f_{n-1} \\ & \cdot & \cdot \\ & I & \cdot \\ 0 & \cdots & 0\ f_0 \end{bmatrix} = \begin{bmatrix} a_{m-n+1} & \cdots & a_{m-1} & e_m \\ a_{m-n+2} & & a_m & 0 \\ \cdot & & \cdot & \cdot \\ a_m & \cdots & a_{m+n-2} & 0 \end{bmatrix}, \tag{5.10}$$

where the last column is determined by equating coefficients of z^i, for $m \le i \le m+n-1$ in (4.49). Thus, $\det(H_{m,n}) \neq 0$ and $\det(f_0) \neq 0$ implies that $\det(H_{m,n-1}) \neq 0$. Thus, (5.2) implies (5.1).

In a similar fashion, by defining

$$U_m^*(z) = E_m^*(z) - zg_0^*r_0^{*-1}P_{m-1}^*(z), \tag{5.11}$$

$$V_n^*(z) = F_{n-1}^*(z) - zg_0^*r_0^{*-1}Q_{n-1}^*(z), \tag{5.12}$$

it can be shown that (5.3) implies (5.1).

To obtain the inverse formula, substitution of (5.6), (5.7) , (5.11) and (5.12), after normalization by (5.5), into equation (4.38) gives

$$H_{m,n}^{-1} = \begin{bmatrix} f_{n-1} & \cdots & f_0 \\ \vdots & & \\ f_0 & & \end{bmatrix} \begin{bmatrix} q_{n-1}^* & \cdots & q_0^* \\ & \ddots & \\ & & q_{n-1}^* \end{bmatrix} - \begin{bmatrix} q_{n-2} & q_0 & 0 \\ \vdots & \ddots & \\ q_0 & & \\ 0 & & \end{bmatrix} \begin{bmatrix} 0 & f_{n-1}^* & \cdots & f_1^* \\ & \ddots & & \vdots \\ & & & f_{n-1}^* \\ & & & 0 \end{bmatrix}$$

$$+ \begin{bmatrix} q_{n-2} & q_0 & 0 \\ \vdots & \ddots & \\ q_0 & & \\ 0 & & \end{bmatrix} (g_0 - g_0^*) \begin{bmatrix} q_{n-1}^* & \cdots & q_0^* \\ & \ddots & \\ & & q_{n-1}^* \end{bmatrix}. \tag{5.13}$$

But, equations (4.49) and (4.50) imply that

$$E_m^*(z)F_{n-1}(z) - F_{n-1}^*(z)E_m(z) = z^{m+n}\left\{F_{n-1}^*(z)G(z) - G^*(z)F_{n-1}(z)\right\}. \tag{5.14}$$

Consequently,

$$F_{n-1}^*(z)G(z) - G^*(z)F_{n-1}(z) = 0 \tag{5.15}$$

and, in particular,

$$g_0 = g_0^*. \tag{5.16}$$

Thus, (5.13) is exactly (5.4), since the last product cancels.

Remark 1: Corollary 5.1 can be proved directly from (4.7), (4.8), (5.1) and (5.2). Indeed, using the same arguments as in Lemma 4.2, we can obtain

$$\begin{bmatrix} r_0^{*-1}Q_{n-1}^*(z) & -r_0^{*-1}P_{n-1}^*(z) \\ -f_0^{*-1}F_{n-1}^*(z) & f_0^{*-1}E_n^*(z) \end{bmatrix} \begin{bmatrix} E_m(z)f_0^{-1} & P_{m-1}(z)r_0^{-1} \\ F_{n-1}(z)f_0^{-1} & Q_{n-1}(z)r_0^{-1} \end{bmatrix} = z^{m+n-1} \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \tag{5.17}$$

and the commutative relationship

$$\begin{bmatrix} E_m(z)f_0^{-1} & P_{m-1}(z)r_0^{-1} \\ F_{n-1}(z)f_0^{-1} & Q_{n-1}(z)r_0^{-1} \end{bmatrix} \begin{bmatrix} r_0^{*-1}Q_{n-1}^*(z) & -r_0^{*-1}P_{m-1}^*(z) \\ -f_0^{*-1}F_{n-1}^*(z) & f_0^{*-1}E_m^*(z) \end{bmatrix} = z^{m+n-1} \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}. \tag{5.18}$$

Consequently, we can normalize our Padé forms according to (5.5) and the formulae will follow in a fashion similar to the proof of Theorem 4.3.

The actual proof, in addition to being simpler, serves to illustrate the existence of Frobenius-type relationships (generalized from the scalar case (c.f., Gragg [26]) to the matrix case) between matrix Padé forms of types (m,n), (m,n-1) and (m-1,n-1). These relationships, which exist under the assumptions of Corollary 5.1, are given by (5.6), (5.7), (5.11) and (5.12) (see also [10] , [14] , [13]).

Remark 2: From (5.17), it follows from equating coefficients of degree m+n-1 that

$$e_m^* q_{n-1} = f_0^* r_0 \tag{5.19}$$

and

$$q_{n-1}^* e_m = r_0^* f_0. \tag{5.20}$$

Thus, if the conditions of Corollary 5.1 are satisfied, then e_m^*, q_{n-1}, q_{n-1}^* and e_m are all nonsingular. Normalizing (4.7), (4.8), (4.49) and (4.50) by setting

$$r_0 = r_0^* = e_m = e_m^* = I, \tag{5.21}$$

rather than by (5.5), we obtain

$$H_{m,n} \cdot \left[q_{n-1}, \cdots, q_0 \right]^t = \left[0, \cdots, 0, I \right]^t \tag{5.22}$$

$$\left[q_{n-1}^*, \cdots, q_0^* \right] \cdot H_{m,n} = \left[0, \cdots, 0, I \right] \tag{5.23}$$

$$H_{m,n} \cdot \left[f_{n-1}, \cdots, f_0 \right]^t = \left[I, 0, \cdots, 0 \right]^t \tag{5.24}$$

$$-\left[f_{n-1}^{*},\cdots,f_{0}^{*}\right]\cdot H_{m,n}=\left[I,0,\cdots,0\right]. \tag{5.25}$$

These conditions, together with the requirement that $\det(q_{n-1})\neq 0$ and $\det(q_{n-1}^{*})\neq 0$, are exactly the conditions given by Gohberg and Heinig [25] in deriving the inverse formula (5.4). Because of the different normalization requirement, their formula includes the term q_{n-1}^{-1} between the first two matrices and q_{n-1}^{*-1} between the last two matrices. This is permissible because of (5.19), (5.20) and (5.21). In the scalar case, this is the well known formula of Gohberg and Semencul [23] .

Remark 3: The assumptions of Corollary 5.1, which are equivalent to conditions (5.22), (5.23), (5.24) and (5.25) of Gohberg and Heinig, are far more restrictive than the assumptions of Theorem 4.3, which are equivalent to (4.45), (4.46), and the block Yule-Walker equations (4.47) and (4.48). The formula of Gohberg and Heinig has the additional requirement that q_{n-1} and q_{n-1}^{*} be nonsingular (which is equivalent to $H_{m,n-1}$ being nonsingular). Thus, for example, (4.38) can be used to obtain the inverse of

$$H_{2,3}=\begin{bmatrix} I & 0 & 0\\ 0 & 0 & I\\ 0 & I & 0\end{bmatrix}, \tag{5.26}$$

whereas, (5.4) cannot be applied.

Remark 4: Since the assumptions of Corollary 5.1 require that not only $H_{m,n}$ but also $H_{m,n-1}$ be nonsingular, it should be possible to express the inverse of $H_{m,n-1}$ in closed form as well. Indeed, by deriving Frobenius-type identities similar to (5.6), (5.7), (5.11) and (5.12) (c.f., Bultheel [10] , [14] , [13]), the matrix Padé form of type (m-1,n-2) can be expressed in terms of matrix Padé forms of type (m,n-1) and (m-1,n-1). Then, substituting the Padé forms of type (m,n-1) and (m-1,n-2) into (4.38) (with n replaced by n-1) and simplifying, we obtain as another corollary to Theorem 4.3 the second inverse formula of Gohberg and Heinig, namely,

$$H_{m,n-1}^{-1}=\begin{bmatrix} f_{n-2} & \cdots & f_0\\ \vdots & \iddots & \\ f_0 & & \end{bmatrix}\begin{bmatrix} q_{n-1}^{*} & \cdots & q_1^{*}\\ & \ddots & \\ & & q_{n-1}^{*}\end{bmatrix}-\begin{bmatrix} q_{n-2} & & q_0\\ \vdots & \iddots & \\ q_0 & & \end{bmatrix}\begin{bmatrix} f_{n-1}^{*} & \cdots & f_1^{*}\\ & \ddots & \\ & & f_{n-1}^{*}\end{bmatrix}. \tag{5.27}$$

Here, we have again normalized according to (5.5). We also note that the Gohberg-Heinig formulae given here are both determined from formula (4.38). Additional formulae , based on (4.39) rather than (4.38), can also be derived.

Remark 5: Gohberg and Heinig prove their formulae with the coefficients over a non-commutative algebra. Our formulae and results also carry over with minor alterations. In particular, Theorem 4.1 and Corollary 5.1 would both require that (4.9) be equivalent to (4.10) and (4.11), simultaneously.

5.3. The Inverse Formulae of Gohberg-Krupnik

Let $(L_{m-2}(z),M_{n-2}(z),N(z))$ and $(L^{*}_{m-2}(z),M^{*}_{n-2}(z),N^{*}(z))$ be a RMPFo and LMPFo, respectively, of type (m-2,n-2) for A(z). These matrix Padé forms then satisfy

$$A(z)M_{n-2}(z)-L_{m-2}(z)=z^{m+n-3}N(z) \tag{5.28}$$

$$M^{*}_{n-2}(z)A(z)-L^{*}_{m-2}(z)=z^{m+n-3}N^{*}(z). \tag{5.29}$$

The inverse of $H_{m,n}$ in terms of matrix Padé forms of types (m-2,n-2) and (m-1,n-1) is given by

Corollary 5.2. Let the matrix Padé forms identified by (4.7), (4.8), (5.28) and (5.29) be given. Then, the following statements are equivalent:

$$det(H_{m,n})\neq 0 \quad \text{and} \quad det(H_{m-1,n-1})\neq 0, \tag{5.30}$$

$$det(n_0)\neq 0,\ det(q_0)\neq 0,\ \text{and}\ det(r_0)\neq 0, \tag{5.31}$$

and

$$det(n^{*}_0)\neq 0,\ det(q^{*}_0)\neq 0,\ \text{and}\ det(r^{*}_0)\neq 0. \tag{5.32}$$

In addition, if any (and therefore all) of the conditions (5.30), (5.31) or (5.32) are satisfied, then

$$H_{m,n}^{-1} = \begin{bmatrix} q_{n-2} & q_0 & 0 \\ \cdot & & \\ & & \\ q_0 & & \\ 0 & & \end{bmatrix} q_0^{-1} \begin{bmatrix} m_{n-2}^* & & m_{-1}^* \\ & & \\ & & \\ & & m_{n-2}^* \end{bmatrix} - \begin{bmatrix} m_{n-3} & & m_0 & 0 & 0 \\ \cdot & & & & \\ m_0 & & & & \\ 0 & & & & \\ 0 & & & & \end{bmatrix} q_0^{*-1} \begin{bmatrix} q_{n-1}^* & & q_0^* \\ & \cdot & \\ & & \\ & & q_{n-1}^* \end{bmatrix}$$

$$+ \begin{bmatrix} q_{n-1} \\ \cdot \\ \cdot \\ \cdot \\ q_0 \end{bmatrix} q_0^{-1} \left[q_{n-1}^*, \cdots, q_0^* \right]. \tag{5.33}$$

Here, the matrix Padé forms have been normalized so that[2]

$$n_0 = n_0^* = r_0 = r_0^* = I \,. \tag{5.34}$$

Proof: To prove that (5.30) is equivalent to (5.31), it follows directly from Theorem 2.3 that $\det(H_{m,n}) \neq 0$ implies that $\det(r_0) \neq 0$ while $\det(H_{m-1,n-1}) \neq 0$ implies that $\det(n_0) \neq 0$ and $\det(q_0) \neq 0$. Conversely, suppose that (5.31) holds. Again, from Theorem 4.1, we have that $\det(n_0) \neq 0$ and $\det(q_0) \neq 0$ implies $\det(H_{m-1,n-1}) \neq 0$. But, then

$$\begin{bmatrix} a_{m-n+1} & \cdots & a_m \\ \cdot & & \cdot \\ \cdot & & \cdot \\ a_m & \cdots & a_{m+n-1} \end{bmatrix} \begin{bmatrix} I & & q_{n-1} \\ \cdot & & \cdot \\ & I & \cdot \\ \cdots & & q_0 \end{bmatrix} = \begin{bmatrix} a_{m-n+1} & \cdots & a_{m-1} & 0 \\ \cdot & & \cdot & \cdot \\ a_{m-1} & \cdots & a_{m+n-2} & 0 \\ a_m & & a_{m+n-1} & r_0 \end{bmatrix}, \tag{5.35}$$

together with the assumption that $\det(r_0) \neq 0$, implies that also $\det(H_{m,n}) \neq 0$.

A similar argument shows that (5.30) is equivalent to (5.32).

To prove (5.33), we first establish some identities. Observe that, under the normalization condition (5.34), $(L_{m-2}(z), M_{n-2}(z), N(z))$, $(L_{m-2}^*(z), M_{n-2}^*(z), N^*(z))$, $(P_{m-1}(z)q_0^{-1}, Q_{n-1}(z)q_0^{-1}, R(z)q_0^{-1})$, and $(q_0^{*-1}P_{m-1}^*(z), q_0^{*-1}Q_{n-1}^*(z), q_0^{*-1}R^*(z))$, satisfy the conditions of Lemma 4.2, with (m,n) replaced by (m-1,n-1). Here, (4.29) becomes

[2] Rather than normalizing with $r_0 = r_0^* = I$, it is equally proper to normalize with $q_0 = q_0^* = I$.

$$\begin{bmatrix} Q_{n-1}(z)q_0^{-1} & M_{n-2}(z) \\ z^2R(z)q_0^{-1} & N(z) \end{bmatrix} \begin{bmatrix} N^*(z) & -M_{n-2}^*(z) \\ -z^2q_0^{*-1}R^*(z) & q_0^{*-1}Q_{n-1}^*(z) \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \tag{5.36}$$

and, in particular,

$$\left\{R(z)q_0^{-1}\right\}N^*(z) = N(z)\left\{q_0^{*-1}R^*(z)\right\}. \tag{5.37}$$

Note that the constant and linear terms in (5.37) yield

$$q_0 = q_0^* \tag{5.38}$$

and

$$q_0^{*-1}(n_1^* - r_1^*) = (n_1 - r_1)q_0^{-1}. \tag{5.39}$$

For later purposes, also observe the identity

$$\begin{bmatrix} q_{n-1} & \cdots & q_0 \\ & & \\ q_0 & & \end{bmatrix} q_0^{-1} \begin{bmatrix} q_{n-1}^* & \cdots & q_0^* \\ & & \\ & & q_{n-1}^* \end{bmatrix} - \begin{bmatrix} q_{n-2} & \cdots & q_0 0 \\ & & \\ q_0 & & \\ 0 & & \end{bmatrix} q_0^{*-1} \begin{bmatrix} 0 q_{n-1}^* & q_0^* \\ & \\ & q_{n-1}^* \end{bmatrix}$$

$$= \begin{bmatrix} q_{n-1} \\ \\ \\ q_0 \end{bmatrix} q_0^{-1} \left[q_{n-1}^*, \cdots, q_0^*\right], \tag{5.40}$$

which follows using (5.38).

Next, we proceed as in Corollary 5.1 by constructing right and left matrix Padé forms of type (m,n) for A(z). Set

$$U_m(z) = \left\{P_{m-1}(z)\left[I + (n_1 - r_1)z\right] - L_{m-2}(z)z^2\right\}q_0^{-1} \tag{5.41}$$

and

$$V_n(z) = \left\{ Q_{n-1}(z)\left[I + (n_1 - r_1)z \right] - M_{n-2}(z)z^2 \right\} q_0^{-1} . \tag{5.42}$$

Then, $U_m(z)$ and $V_n(z)$ provide a RMPFo of type (m,n) for A(z). To see this, note that the degree requirements are clearly satisfied. In addition, the columns of $V_n(z)$ are linearly independent since, in (5.42), $v_0 = I$. Finally,

$$\begin{aligned}
&A(z)V_n(z) - U_m(z) \\
&= \left\{ \left[A(z)Q_{n-1}(z) - P_{m-1}(z) \right] \left[I + (n_1 - r_1)z \right] - z^2 \left[A(z)M_{n-2}(z) - L_{m-2}(z) \right] \right\} q_0^{-1} \\
&= \left\{ z^{m+n-1} R(z) \left[I + (n_1 - r_1)z \right] - z^{m+n-1} N(z) \right\} q_0^{-1} \\
&= z^{m+n-1} \left\{ (r_0 - n_0) + (r_1 + r_0(n_1 - r_1) - n_1)z + z^2 \left[\cdots \right] \right\} q_0^{-1} \\
&= z^{m+n+1} \left\{ \cdots \right\} q_0^{-1} ,
\end{aligned} \tag{5.43}$$

since $n_0 = r_0 = I$.

Similarly, it can be shown that

$$U_m^*(z) = q_0^{*-1} \left\{ \left[I + (n_1^* - r_1^*)z \right] P_{m-1}^*(z) - L_{m-2}^*(z)z^2 \right\} \tag{5.44}$$

$$V_n^*(z) = q_0^{*-1} \left\{ \left[I + (n_1^* - r_1^*)z \right] Q_{n-1}^*(z) - M_{n-2}^*(z)z^2 \right\} . \tag{5.45}$$

provides a LMPFo of type (m,n) for A(z).

Note that (5.42) and (5.45), respectively, yield

$$-\begin{bmatrix} v_{n-1} & \cdots & v_0 \\ & & \\ v_0 & & \end{bmatrix} = \begin{bmatrix} q_{n-1} & \cdots & q_0 \\ & & \\ q_0 & & \end{bmatrix} q_0^{-1} + \begin{bmatrix} q_{n-2} & \cdots & q_0 0 \\ & & \\ q_0 & & \\ 0 & & \end{bmatrix} (n_1 - r_1) q_0^{-1}$$

$$- \begin{bmatrix} m_{n-3} & \cdots & m_0 0 0 \\ & & \\ m_0 & & \\ 0 & & \\ 0 & & \end{bmatrix} q_0^{-1} \qquad (5.46)$$

and

$$\begin{bmatrix} v_n^* & \cdots & v_1^* \\ & & \\ & & v_n^* \end{bmatrix} = q_0^{*-1} \begin{bmatrix} 0 q_{n-1}^* & q_1^* \\ & \\ & q_{n-1}^* \\ & 0 \end{bmatrix} + q_0^{*-1}(n_1^* - r_1^*) \begin{bmatrix} q_{n-1}^* & q_0^* \\ & \\ & q_{n-1}^* \end{bmatrix}$$

$$- q_0^{*-1} \begin{bmatrix} m_{n-2}^* & m_{-1}^* \\ & \\ & m_{n-2}^* \end{bmatrix} \qquad (5.47)$$

where $m_{-1}^* = 0$. Substituting (5.46) and (5.47) into (4.38), and rearranging terms, we obtain

$$H_{m,n}^{-1} = \begin{bmatrix} q_{n-2} & q_0 0 \\ & \\ q_0 & \\ 0 & \end{bmatrix} q_0^{-1} \begin{bmatrix} m_{n-2}^* & m_{-1}^* \\ & \\ & m_{n-2}^* \end{bmatrix} - \begin{bmatrix} m_{n-3} & m_0 0 0 \\ & \\ m_0 & \\ 0 & \\ 0 & \end{bmatrix} q_0^{*-1} \begin{bmatrix} q_{n-1}^* & q_0^* \\ & \\ & q_{n-1}^* \end{bmatrix}$$

$$+ \begin{bmatrix} q_{n-1} & \cdots & q_0 \\ & & \\ q_0 & & \end{bmatrix} q_0^{-1} \begin{bmatrix} q_{n-1}^* & \cdots & q_0^* \\ & & \\ & & q_{n-1}^* \end{bmatrix} - \begin{bmatrix} q_{n-2} & \cdots & q_0 0 \\ & & \\ q_0 & & \\ 0 & & \end{bmatrix} q_0^{*-1} \begin{bmatrix} 0 q_{n-1}^* & q_1^* \\ & \\ & q_{n-1}^* \\ & 0 \end{bmatrix}$$

$$+\begin{bmatrix} q_{n-2} & \cdots & q_0 & 0 \\ & & & \\ q_0 & & & \\ 0 & & & \end{bmatrix}\left\{(n_1-r_1)q_0^{-1}-q_0^{*-1}(n_1^*-r_1^*)\right\}\begin{bmatrix} q_{n-1}^* & \cdots & q_0^* \\ & & \\ & & q_{n-1}^* \end{bmatrix}. \tag{5.48}$$

But, using (5.38) and (5.39), it is easy to see that (5.48) is exactly (5.33)

Remark 1. The inverse formula (5.33) can also be determined by bordering techniques. Indeed, equation (5.35) can be further manipulated to obtain

$$H_{m,n}^{-1}=\begin{bmatrix} & H_{m-1,n-1}^{-1} & 0 \\ & & \vdots \\ 0 & \cdots & 0 \end{bmatrix}+\begin{bmatrix} q_{n-1} \\ \vdots \\ q_0 \end{bmatrix}q_0^{-1}\begin{bmatrix} q_{n-1}^* & \cdots & q_0^* \end{bmatrix}. \tag{5.49}$$

Equation (4.38) applied to $H_{m-1,n-1}^{-1}$, along with simplification using Lemma 4.2, converts (5.49) into (5.33).

The present proof takes its cue from the approach of sections 4 and 5. In each case, the inverse formula is obtained from (4.38) using Frobenius-type identities for matrix Padé forms. The Frobenius-type identities (5.41), (5.42), (5.44) and (5.45) used here can be found in [13] (see also chapter 2).

Remark 2. Notice that, if the matrix Padé forms (4.7), (4.8), (5.28) and (5.29) satisfy the conditions of Corollary 5.2 and are normalized according to (5.34), then

$$H_{m,n}\cdot\left\{\begin{bmatrix} m_{n-2} \\ \\ \\ m_0 \\ 0 \end{bmatrix}-\begin{bmatrix} q_{n-1} \\ \\ \\ q_0 \end{bmatrix}n_1\right\}=\begin{bmatrix} 0 \\ \\ 0 \\ I \\ n_1 \end{bmatrix}-\begin{bmatrix} 0 \\ \\ 0 \\ I \end{bmatrix}n_1=\begin{bmatrix} 0 \\ \vdots \\ 0 \\ I \\ 0 \end{bmatrix}. \tag{5.50}$$

Thus, the second last column of $H_{m,n}^{-1}$ is a combination of the coefficients of $M_{n-2}(z)$ and $Q_{n-1}(z)$. Similarly, one can obtain the second last row of $H_{m,n}^{-1}$ as a combination of the coefficients of $M_{n-2}^*(z)$ and

$Q^*_{n-1}(z)$.

Conversely, suppose $X = [\,x_{n-1}, \cdots, x_0\,]^t$ and $Q = [\,q_{n-1}, \cdots, q_0\,]^t$, respectively, represent the second last and last block columns of the inverse of $H_{m,n}$. Then, if $\det(q_0) \neq 0$, we have that

$$H_{m,n}\left\{\begin{bmatrix} x_{n-1} \\ \\ x_0 \end{bmatrix} - \begin{bmatrix} q_{n-1} \\ \\ q_0 \end{bmatrix} q_0^{-1} x_0\right\} = \begin{bmatrix} 0 \\ \\ 0 \\ I \\ -q_0^{-1}x_0 \end{bmatrix} \tag{5.51}$$

so that

$$H_{m-1,n-1}\left\{\begin{bmatrix} x_{n-1} \\ \\ x_1 \end{bmatrix} - \begin{bmatrix} q_{n-1} \\ \\ q_1 \end{bmatrix} q_0^{-1} x_0\right\} = \begin{bmatrix} 0 \\ 0 \\ I \end{bmatrix}. \tag{5.52}$$

This implies that

$$M_{n-2}(z) = z^{-1}\left\{X(z) - Q_{n-1}(z) q_0^{-1} x_0\right\} \tag{5.53}$$

is a RMPFo denominator of type (m-2,n-2) for A(z). Similarly, we can obtain a LMPFo denominator of type (m-2,n-2) when we have the last and second last block rows of the inverse of $H_{m,n}$. Then substitution into (5.33) yields

$$H^{-1}_{m,n} = \begin{bmatrix} q_{n-2} & q_0 & 0 \\ \cdot & & \\ & & \\ q_0 & & \\ 0 & & \end{bmatrix} q_0^{-1} \begin{bmatrix} x^*_{n-1} & x^*_0 \\ & \\ & x^*_{n-1} \end{bmatrix} - \begin{bmatrix} x_{n-2} & x_0 & 0 \\ \cdot & & \\ & & \\ x_0 & & \\ 0 & & \end{bmatrix} q_0^{*-1} \begin{bmatrix} q^*_{n-1} & q^*_0 \\ & \\ & q^*_{n-1} \end{bmatrix}$$

$$+ \begin{bmatrix} q_{n-2} & q_0 & 0 \\ \cdot & & \\ & & \\ q_0 & & \\ 0 & & \end{bmatrix} q_0^{-1}(x^*_0 - x_0) q_0^{*-1} \begin{bmatrix} q^*_{n-1} & q^*_0 \\ & \\ & q^*_{n-1} \end{bmatrix} + \begin{bmatrix} q_{n-1} \\ \cdot \\ \cdot \\ \cdot \\ q_0 \end{bmatrix} q_0^{-1} \left[q^*_{n-1}, \cdots, q^*_0\right]; \tag{5.54}$$

whereas, substitution into (4.38) gives

$$H_{m-1,n-1}^{-1} = \begin{bmatrix} q_{n-2} & q_0 \\ \cdot & \cdot \\ q_0 & \end{bmatrix} q_0^{-1} \begin{bmatrix} x_{n-1}^* & x_1^* \\ & \cdot \\ & x_{n-1}^* \end{bmatrix} - \begin{bmatrix} x_{n-2} & x_0 \\ \cdot & \cdot \\ x_0 & \end{bmatrix} q_0^{*-1} \begin{bmatrix} q_{n-1}^* & q_1^* \\ & \cdot \\ & q_{n-1}^* \end{bmatrix}. \tag{5.55}$$

Remark 3. In the scalar case, if $X = [\, x_{n-1}, \cdots, x_0 \,]^t$, and $Q = [\, q_{n-1}, \cdots, q_0 \,]^t$, respectively, represent the second last and last columns of the inverse of $H_{m,n}$, and $q_0 \neq 0$, then (5.54) and (5.55) reduce to

$$H_{m,n}^{-1} = q_0^{-1} \left\{ \begin{bmatrix} q_{n-2} & q_0 & 0 \\ \cdot & \cdot & \\ q_0 & & \\ 0 & & \end{bmatrix} \begin{bmatrix} x_{n-1} & x_0 \\ & \cdot \\ & x_{n-1} \end{bmatrix} - \begin{bmatrix} x_{n-2} & x_0 & 0 \\ \cdot & \cdot & \\ x_0 & & \\ 0 & & \end{bmatrix} \begin{bmatrix} q_{n-1} & q_0 \\ & \cdot \\ & q_{n-1} \end{bmatrix} \right.$$

$$\left. + \begin{bmatrix} q_{n-1}^2 & \cdots & q_{n-1}q_0 \\ & & \\ q_{n-1}q_0 & \cdots & q_0^2 \end{bmatrix} \right\}. \tag{5.56}$$

and

$$H_{m-1,n-1}^{-1} = q_0^{-1} \left\{ \begin{bmatrix} q_{n-2} & q_0 \\ \cdot & \cdot \\ q_0 & \end{bmatrix} \begin{bmatrix} x_{n-1} & x_1 \\ & \cdot \\ & x_{n-1} \end{bmatrix} - \begin{bmatrix} x_{n-2} & x_0 \\ \cdot & \cdot \\ x_0 & \end{bmatrix} \begin{bmatrix} q_{n-1} & q_1 \\ & \cdot \\ & q_{n-1} \end{bmatrix} \right\} \tag{5.57}$$

These are the original formulae of Gohberg and Krupnik [24] .

Remark 4. Following the approach of section 4, we can also obtain conditions and inverse formulae for $H_{m-1,n-1}$ and $H_{m,n}$ when the first and second block column, along with the first and second block row, of the inverse of $H_{m,n}$ is given (c.f., Iohvidov [29]). Here, conditions and inverse formulae for $H_{m-1,n-1}$ and $H_{m,n}$ are stated in terms of matrix Padé forms of type (m,n-1) and (m+1,n-2). Additional formulae,

based on (4.39) rather than (4.38) can also be given.

5.4. Additional Inverse Formulae

The Frobenius-type relationships given in this chapter are but a small sample of similar recurrence relationships which exist between matrix Padé forms and which are developed by Bultheel ([10] , [14] , [13]). All the relationships require the existence of inverses of certain coefficients in the Padé forms involved. These requirements are always satisfied for normal matrix power series (where $H_{m,n}$ is nonsingular for all m and n). For this restricted class of power series, many of the recursive relationships provide directly algorithms for the computation of Padé forms. Depending on the path (within the Padé table) determined by the recurrence, Bultheel observes that the algorithms ([1] , [7] , [21] , [33] , [36] , [41] , [47]) which explicitly or implicitly compute the inverse of Hankel or Toeplitz matrices are equivalent to using an appropriate recurrence formula.

For a subset of these relationships, this chapter shows that each recurrence yields a separate closed formula for the inverse of a block Hankel. Algorithms based on recurrences which specify computations along an off-diagonal path (e.g., [1] , [7] , [41] [47]) yield closed formulae expressed by (4.38), (4.39) and (5.33). Those that specify computations along a staircase (e.g., [21] , [33] , [36]) yield formulae (5.4) and (5.27); whereas, those that specify computations along an anti-diagonal path yield (4.56) and (4.57). Additional closed formulae can be derived corresponding to other recurrences given by Bultheel.

5.5. Computation of Block Hankel Inverses

Formulas (5.4), (5.27), and (5.33) are equivalent to ones given by Gohberg-Heinig, and Gohberg-Krupnik, whereas (4.38), (4.39), (4.56) and (4.57) are new. A major advantage of the new formulae is that the underlying assumptions are far less restrictive than they are for (5.4), (5.27) and (5.33). Whereas, the new formulae require only that $H_{m,n}$ be nonsingular, the latter also require that an additional submatrix be

nonsingular. In addition, necessary and sufficient conditions for the existence of $H_{m,n}^{-1}$ are directly available from the coefficients of Padé forms. This provides a significant computational advantage.

Relaxed conditions provide little computational gain, however, if the available algorithms can function only under the more severe restrictions of normality. Unfortunately, this is true for most algorithms that compute nonscalar Padé forms or decompose block Hankel matrices. One exception in this regard is the MPADE algorithm of chapter 3. This algorithm is based on a recurrence relationship between Padé forms at successive nonsingular nodes along an off-diagonal path of the matrix Padé table (or, by reversing coefficients, along an anti-diagonal path). When the power series is normal, or, less restrictively, when all minors of the associated Hankel matrix are nonsingular (e.g., when the block Hankel matrix is positive definite), all the nodes along the path are nonsingular and then the recurrence relationship reduces to (5.42), which is one of many given by Bultheel. The methods based on this relationship are special cases of the MPADE algorithm.

For purposes of expressing the inverse of $H_{m,n}$ in terms of the new formulae (4.38), (4.39), (4.56) and (4.57), the MPADE algorithm is particularly suitable. Singularity is detected with no additional effort. When $H_{m,n}$ is nonsingular, the necessary Padé forms (i.e., the solutions of the associated block Yule-Walker equations) appearing in the formulae are simultaneously available on termination. The algorithm has no restrictions of normality. In addition, intermediate results enable the computation of the inverses of any nonsingular principal minors. When the power series is normal, the cost is the same as that of previously mentioned algorithms.

Using fast polynomial arithmetic in the normal case, Bitmead and Anderson [6] indicate that their scalar algorithm, based on a divide-and-conquer partitioning of the Hankel matrix, can be generalized to the nonscalar case with a cost complexity of $O(p^3 n \log^2 n)$. Under somewhat relaxed normality conditions (i.e., near-normality), Labahn [32] also gives an algorithm, an adaption of MPADE, with the same complexity.

In the scalar case, one call of an algorithm given by Cabay and Choi [16] can be used to construct the inverse formulae (4.38), (4.39), (4.56) or (4.57) with cost complexity $O(n \log^2 n)$ under no restrictions of normality. There are other methods (c.f. Sugiyama [44] for a survey) that also have the same complexity and succeed in the abnormal case. In particular, this is true of the scalar method of Brent et al [8] . They use two calls of a fast anti-diagonal GCD algorithm, EMGCD, to determine the two Padé forms required by the Gohberg-Semencul formula (5.4). The algorithm succeeds immediately if both $H_{n,n}$ and $H_{n,n-1}$ are nonsingular. If $H_{n,n-1}$ is singular (but $H_{n,n}$ is not), then a nonsingular matrix $H_{n,n+1}$ is first constructed. Two additional calls of the anti-diagonal algorithm are then made to yield the two Padé forms required by the second formula (5.27) of Gohberg and Semencul. By computing the inverse of $H_{n,n}$ using formula (4.56) or (4.57) their algorithm can now be altered so as to only require one call of their anti-diagonal algorithm.

Another significant drawback in basing an algorithm on the inverse formulae of Gohberg and Semencul has been observed by Bunch [15] . Because two distinct formulae are used to represent $H_{n,n}^{-1}$, depending on the status of singularities of relevant minors, the results are subject to numerical instability. No such problems are encountered with the use of the new formulae.

Chapter 6

Summary and Suggestions for Future Research

We have considered the problem of determining an adequate definition for a rational approximant of a formal matrix power series and also, given a suitable definition, the problem of computing it. We have restricted our attention to square matrix power series.

In attempting to extend the notion of Padé approximation to matrix power series, we have followed the classical theory of Padé approximants for scalar power series. We introduce the notion of a matrix Padé form, which always exist but may not be unique, and also the notion of matrix Padé fraction, which is unique but need not exist. The definition of matrix Padé form is meant to be as broad as possible. By constructing all the matrix Padé forms of type (m,n), it is always possible to determine ones for which the denominator is invertible, should one exist.

The notion of a matrix power series remainder sequence introduced in chapter 3 is a generalization of one given by Cabay and Kossowski [17] for scalar power series. The cofactor sequence, which is shown to be associated with the remainder sequence, yields directly all the matrix Padé fractions at the nonsingular nodes of a particular off-diagonal path of the matrix Padé table. By determining also the (unique) matrix Padé form at nodes preceeding the nonsingular nodes, we are able to compute matrix Padé fractions iteratively from one nonsingular node to the next. The resulting algorithm, MPADE, is at least as fast as other algorithms for computing matrix Padé fractions, and it is the only one that succeeds in the abnormal case.

The relationship between matrix Padé forms and block Hankel matrices has also between investigated in the study of the invertibility of such matrices. We show that a necessary and sufficient condition for the nonsingularity of a block Hankel matrix is the existence of four matrix Padé forms satisfying special conditions. Furthermore, when a block Hankel matrix is nonsingular we have presented a new set of closed formulae for the inverse in terms of these four matrix Padé forms. Unlike other inverse formulae, no extra conditions are required for these formulae to hold. The resulting formulae are practical since the required matrix Padé forms can be calculated by the MPADE algorithm of chapter 3. Indeed, the iterative

algorithm also calculates inverses of all nonsingular principal minors. It is at least as fast as other algorithms for computing inverses in the block case and is the only one that imposes no additional conditions. In the scalar case the required Padé forms can be calculated by the off-diagonal algorithm of Cabay and Choi [16] . The result is also an iterative algorithm that calculates inverses of some of the nonsingular principal minors and is faster than existing algorithms.

By using various Frobenius-type identities for matrix Padé forms, the formulae of chapter 4 give additional inverse formulae. Specifically, we have shown that the well known, more restrictive formulae of Gohberg-Krupnik [24] , Gohberg-Semencul [23] and Gohberg-Heinig [25] follow from our closed formulae. Thus our approach, using matrix Padé forms, results in the unification of existing work on Hankel inverses.

With regards to the study of matrix Padé forms, there are many areas that can be researched. The MPADE algorithm can be improved in a number of ways. We expect that the cost of the decomposition of $T_{(\mu-1),\mu}$ in step 2.7 and consequently of $S_{(\mu-1),\mu}$ in step 3.1, can be achieved in time faster than $O(p^3 s_i^3)$, by taking advantage of the special structure of Sylvester matrices. The algorithm would also experience an improvement if it were possible to identify additional points between nonsingular nodes for which Theorem 3.5 is valid. This would improve the algorithm by decreasing the s_i. This, and in general the nature of matrix Padé forms between nonsingular nodes, is a subject for further research. Finally, by appealing to fast methods for polynomials, it is of interest to attempt to develop a recursive divide-and-conquer version of MPADE, in a fashion similar to one developed by Cabay and Choi [16] for the scalar case. This was done in the matrix case by Labahn [32] but under the restriction of near-normality.

For normal and nearly-normal power series, progressing from one nonsingular node to the next is equivalent to power series division of the residuals associated with the nonsingular nodes (because $S_{(\mu-1),\mu}$ in step 3.1 of MPADE, with the exception of one column, reduces to a triangular matrix). Thus, in this case when, in addition, A(z) and B(z) are matrix polynomials, there is a strong analogy between MPADE and Euclid's algorithm (for the scalar case, see Cabay and Choi [16]). It is thus of interest to use MPADE to

calculate the greatest common divisor of two matrix polynomials.

The calculation of a greatest common divisors of two matrix polynomials plays an important role in the study of minimal partial realizations of matrix sequences (c.f., Gragg and Lindquist [27]). The calculation problems in the matrix case, however, are significantly more complicated than the scalar case. Polynomial division is not always possible, hence Euclid's algorithm is not always applicable. Furthermore, the notion of relative primeness differs from the scalar case. Two matrix polynomials are relatively prime if their only common divisor is a unimodular polynomial, i.e., a matrix polynomial having a constant determinant (and hence, having a matrix polynomial as an inverse). Unlike the scalar case , trivial divisors (i.e., unimodular polynomials) may have nonzero degree. Factoring out a common divisor may result in a rational expression having higher degrees than before (although decreasing the degrees of the determinants).

At present we have partial results when matrix Padé forms are used in the study of the matrix GCD problem. For example, the numerator and denominater of a matrix Padé fraction occurring at a nonsingular node are relatively prime. The only common factors of the numerator and denominater of a predecessor Padé form must have a determinant of the form z^k. If the (m,n) node of the Padé table of (A(z),B(z)) is nonsingular, where A(z) and B(z) have degrees n and m, respectively, then A(z) and B(z) are relatively prime.

The closed inverse formulae presented opens up many new potential research topics. At present we are also investigating closed inverse formulae for Sylvester matrices of the form (2.21) where it is not assumed that b_0 is nonsingular. More generally, we can follow the approach of chapter 4 to obtain closed inverse formulae for matrices of the form

$$S_{m,n} = \left[\begin{array}{ccc|ccc} a_0 & \cdots & a_n & b_0 & \cdots & b_m \\ & & & & & \\ & & & & & \\ & & & & & \\ a_{m+n+1} & \cdots & a_{m+2n+1} & b_{m+n+1} & \cdots & b_{2m+n+1} \end{array}\right] \tag{6.1}$$

that is, of two Hankel-like matrices adjoined together to form one square matrix. Besides their use in Padé

approximantion, these matrices are also important in applying the Sylvester criterion for the relative primeness of two polynomials and for studying propagation through layered media . Necessary and sufficient conditions for determination of the nonsingularity of such generalized Sylvester matrices in the scalar case depends on the existence of three Hermite-Padé approximants (c.f., Baker [4]) for the pair $(A(z), z^{m-n}B(z))$. A closed formula for the inverse is then given in terms of these Hermite-Padé approximants. Unlike existing methods (c.f., Kailath et al [31]), our formula requires only that $S_{m,n}$ is nonsingular. The approach taken essentially follows the methods introduced in chapter 4. At present the formula has many favorable properties, especially in terms of storage requirements and evaluation costs. However, the formula is long and unwieldy so is not yet fully understood. In addition, there remains the problem of obtaining an efficient algorithm for the calculation of the required Hermite-Padé approximants that avoids unnecessary requirements of nonsingularity of certain submatrices. Existing recurrence relationships (c.f., Paszkowski [38]) have strong invertibilty requirements. A recurrence relation that does not require extra invertibility has been obtained and an algorithm, using a similar approach found with the MPADE algorithm , is in progress. It is also expected that corresponding formulae, similar to the scalar formulae, also exists in the block generalized Sylvester case.

There is also the problem of obtaining a closed inverse formula for square matrices of the form

$$H'_{n_1,\cdots,n_k} = \begin{bmatrix} a_0 & \cdots & a_{n_1} & | & & | & b_0 & \cdots & b_{n_k} \\ & & & | & \cdots & | & & & \\ & & & | & & | & & & \\ a_u & \cdots & a_{u+n_1} & | & & | & b_u & \cdots & b_{u+n_k} \end{bmatrix} \tag{6.2}$$

where $u = n_1 + \cdots + n_k + k - 1$, that is, in the case where the matrix consists of k Hankel-like matrices where k is arbitrary. This is a problem in both the scalar and block cases. Again it is expected that necessary and sufficient conditions for the inverse will depend on the existence of certain Hermite-Padé approximants (or block Hermite-Padé approximants), and that a closed formula can be given in terms of these approximants. We note that this problem ranges from the very structured case (k = 1) to the totally unstructured (k = u) case.

An interesting corollary of any inverse formulae in the k Hankel-like case is that it will then provide inverse formulae in the block Hankel matrix case where the blocks are no longer square matrices, but rather matrices of a specific rectangular size. This follows from the ability to transform these block matrices into other block k Hankel-like matrices. For example, if

$$H_n = \begin{bmatrix} a_0 & \cdots & a_n \\ & & \\ a_{pn} & \cdots & a_{(p+1)n} \end{bmatrix} \tag{6.3}$$

where the a_i are 1 x p vectors

$$a_i = (b_i, \ldots, c_i), \tag{6.3}$$

then H_n can be transformed into

$$H'_{n,\ldots,n} = \left[\begin{array}{ccc|c|ccc} b_0 & \cdots & b_n & & c_0 & \cdots & c_n \\ & & & \cdots & & & \\ b_{pn} & \cdots & b_{(p+1)n} & & c_{pn} & \cdots & c_{(p+1)n} \end{array}\right] \tag{6.4}$$

by a permutation of the columns. H_n is nonsingular if and only if $H'_{n,\ldots,n}$ is nonsingular. In addition, any closed inverse formula for the scalar matrix $H'_{n,\ldots,n}$ will provide a closed inverse formula for H_n. When the a_i are q x p matrices, with $q > 1$, then H_n will be transformed into a block matrix $H'_{n,\ldots,n}$ and so a closed formula for the block matrix will provide an inverse formula for H_n in this case.

Chapter 7

Appendix: Validity of Initialization in MPADE Algorithm

It is required to show, in the case where $m < n$, that the $(U_1(z),V_1(z))$ computed by MPADE is a RMPFr for (A(z),B(z)) at the first nonsingular node (m_1,n_1) along the m-n offdiagonal path of the right matrix Padé table for (A(z),B(z)). In addition, we need to show that $(P_1(z),Q_1(z))$ is the predecessor of this nonsingular node.

To show the validity of the initialization step when $m < n$, we need

Lemma 7.1. (U(z),V(z)) is an RMPFo (RMPFr) of type (m,n) for (A(z),B(z)) if and only if $(z^k U(z),V(z))$ is an RMPFo (RMPFr) of type (m+k,n) for $(z^k A(z),B(z))$

Proof: The result follows directly from (2.11).

When $m < n$, one complete pass through MPADE is required before $(U_1(z),V_1(z))$ are computed. Given the initial values (m_0,n_0), $(U_0(z),V_0(z))$ and $(P_0(z),Q_0(z))$ defined by steps 1.6 and 1.7 of MPADE (note that (m_0,n_0) defined here is different from the initialization given in (3.2)), the residuals computed in steps 2.5 and 2.6 then satisfy

$$R_0(z) = B(z) \tag{7.1}$$

and

$$W_0(z) = z^{n-m-1} \cdot A(z), \tag{7.2}$$

where, for convenience, the modulo operation has been dropped. Let s_0 be the smallest positive integer in step 2.7 such that $det(T'_{(s_0-1),s_0}) \neq 0$, where $T'_{(s_0-1),s_0}$ is determined from $R_0(z)$ and $W_0(z)$. Clearly, because of (7.2), $s_0 \geq n-m$.

Let $(U'(z),V'(z))$ be the RMPFr of type (s_0-1,s_0) for $(W_0(z),R_0(z))$ computed in step 3.1. Then, from (7.1) and (7.2) and Lemma 7.1, it follows that $U'(z)$ can be expressed as

$$U'(z) = z^{n-m-1} \cdot U^*(z), \tag{7.3}$$

where $(U^*(z),V'(z))$ is the RMPFr of type (s_0+m-n, s_0) for (A(z),B(z)) at the first nonsingular node along the m-n offdiagonal path of the right matrix Padé table for (A(z),B(z)).

Thus, it is required to show that $(U_1(z),V_1(z))$ computed by MPADE is equal to $(U^*(z),V'(z))$. But, step 5.3 yields

$$\begin{bmatrix} U_1(z) \\ V_1(z) \end{bmatrix} = \begin{bmatrix} U_0(z) & P_0(z) \\ V_0(z) & Q_0(z) \end{bmatrix} \cdot \begin{bmatrix} I & 0 \\ 0 & z^2 \cdot I \end{bmatrix} \cdot \begin{bmatrix} V'(z) \\ U'(z) \end{bmatrix}$$

$$= \begin{bmatrix} 0 & z^{m-n-1} \cdot I \\ I & 0 \end{bmatrix} \cdot \begin{bmatrix} I & 0 \\ 0 & z^2 \cdot I \end{bmatrix} \cdot \begin{bmatrix} V'(z) \\ z^{n-m-1} \cdot U^*(z) \end{bmatrix}$$

$$= \begin{bmatrix} U^*(z) \\ V'(z) \end{bmatrix} \tag{7.4}$$

as required. In addition, steps 5.1 and 5.2 yield

$$m_1 = m_0+s_0 = s_0+m-n \tag{7.5}$$

and

$$n_1 = n_0+s_0 = s_0, \tag{7.6}$$

which specifies correctly the first nonsigular node along the m-n offdiagonal.

In a similar fashion, it can be shown that $(P_1(z),Q_1(z))$ computed in step 5.3 is the predecessor of $(U_1(z),V_1(z))$.

References

[1] H. Akaike , "Block Toeplitz Matrix Inversion ", *SIAM J. Appl. Math* , Vol. 24, 1973, pp. 234-241 .

[2] G.A. Baker, *Essentials of Padé Approximants*, Academic Press, New York, 1975.

[3] G.A. Baker and P. Graves-Morris, *Padé Approximants Part I: Basic Theory*, Addison-Wesley , Reading , 1981.

[4] G.A. Baker, "Invariance Properties in Hermite-Padé Approximation Theory", *J. of Comp. and Appl. Math*, Vol. 11, No. 1, 1984, pp. 49-56.

[5] F.L. Bauer, The Quotient-Difference and Epsilon Algorithms, in *On Numerical Approximation*, R.E. Langer (ed.), University of Wisconsin Press, Madison, 1959, pp. 361-370.

[6] R.R. Bitmead and B.D.O. Anderson , "Asmptotically Fast Solutions of Toeplitz and Related Systems of Linear Equations", *Linear Algebra Appl.* , Vol. 34 , 1980, pp. 103-116 .

[7] N.K. Bose and S. Basu, "Theory and Recursive Computation of 1-D Matrix Padé Approximants", *IEEE Trans. on Circuits and Systems*, Vol. 4, 1980, pp. 323-325 .

[8] R. Brent, F.G. Gustavson and D.Y.Y. Yun, "Fast Solution of Toeplitz Systems of Equations and Computation of Padé Approximants ", *J. of Algorithms* , Vol. 1, 1980, pp. 259-295.

[9] F. Brophy and A.C. Salazar, "Considerations of the Padé Approximant Technique in the Synthesis of Recursive Digital Filters", *IEEE Trans. Audio and Electroacoustics*, Vol. AU-21, No. 6, Dec 1973, pp. 500-505.

[10] A. Bultheel, "Recursive Algorithms for the Matrix Padé Table ", *Math. of Computation*, Vol. 35, 1980 , pp. 875-892.

[11] A. Bultheel , "Recursive Algorithms for Nonnormal Padé Tables", *SIAM J. Appl. Math*, Vol. 39, No. 1, 1980, pp. 106-118.

[12] A. Bultheel, "Algorithms To Compute The Reflection Coefficients of Digital Filters", *Num. Methods of Approx. Theory*, Vol. 7, 1983, .

[13] A. Bultheel , "Recursive Relations for Block Hankel and Toeplitz Systems Part I: Direct Recursions ", *J. Comp. and Appl. Math*, Vol. 10 , 1984, pp. 301-328 .

[14] A. Bultheel , "Recursive Relations for Block Hankel and Toeplitz Systems Part II: Dual Recursions ", *J. Comp. and Appl. Math*, Vol. 10 , 1984, pp. 329-354 .

[15] J.R. Bunch , "Stability of Methods for Solving Toeplitz Systems of Equations", *SIAM J. Sci. Comput.*, Vol. 6, 1985, pp. 349-364 .

[16] S. Cabay and D.K. Choi, Algebraic Computations of Scaled Padé Fractions, *SIAM J. of Computation* , Feb, 1986 , pp. 243-270.

[17] S. Cabay and P. Kossowski , Power Series Remainder Sequences and Padé Fractions over Integral Domains , *J. of Symbolic Computation* , to appear .

[18] D.K. Choi , *Algebraic Compuations of Scaled Padé Fractions* , Ph.D. Thesis, University of Alberta, 1984.

[19] J.E. Dennis , J.F. Traub and R.P. Weber , "The Algebraic Theory of Matrix Polynomials ", *SIAM J. Numer. Anal.*, Vol. 13 , 1976, pp. 831-845.

[20] A. Draux , "The Epsilon Algorithm in a Non-commutative Algebra ", *J. Comp. Appl. Math* , Vol. 19 , 1987, pp. 9-21.

[21] J. Durbin , "The Fitting of Time-Series Models ", *Rev. Inst. Internat. Stat.* , Vol. 28 , 1960 , pp. 233-244 .

[22] O.I. Elgerd, *Control Systems Theory*, McGraw-Hill, New York, 1967.

[23] I.C. Gohberg and A.A. Semencul , "On the Inversion of Finite Toeplitz Matrices and their Continuous Analogs ", *Mat. Issled* , Vol. 2 , 1972, pp. 201-233. (in Russian) .

[24] I.C. Gohberg and N.Ya. Krupnik , "A Formula for the Inversion of Finite Toeplitz Matrices ", *Mat. Issled* , Vol. 2 , 1972, pp. 272-283. (in Russian) .

[25] I.C. Gohberg and G. Heinig , "Inversion of Finite Toeplitz Matrices made of elements of a Non-commutative Algebra ", *Rev. Roum. Math. Pures Et Appl.*, Vol. XIX, No. 5 , 1974, pp. 623-663. (in Russian) .

[26] W.B. Gragg , "The Padé Table and its Relation to Certain Algorithms of Numerical Analysis ", *SIAM.Rev.*, Vol. 14, 1972 , pp. 1-61.

[27] W.B. Gragg and A. Lindquist, "On the Partial Realization Problem ", *Lin. Alg. and Appl.* , Vol. 50 , 1983, pp. 277-319 .

[28] F. de Hoog, "A New Algorithm for Solving Toeplitz Systems of Equations", *Linear Algebra Appl.* , Vol. 88 , 1987, pp. 123-138 .

[29] I.S. Iohvidov , *Hankel and Toeplitz Matrices and Forms* , Birkhauser , Boston, 1982.

[30] T. Kailath , A. Vieira and M. Morf , "Inverses of Toeplitz Operators, Innovations and Orthogonal Polynomials ", *SIAM Review* , Vol. 20, 1978 , pp. 106-119.

[31] T. Kailath , S.Y. Kung and M. Morf , "Displacement Ranks of Matrices and Linear Equations", *J. Math. Anal. Appl.* , Vol. 68 , 1979 , pp. 395-407 .

[32] G. Labahn , *Matrix Padé Approximants* , M.Sc. Thesis, Dep't of Computing Science, University of Alberta, Edmonton, Canada, 1986.

[33] N. Levinson , "The Wiener RMS (root mean square) Error in Filter Design ", *J. Math. and Phys.* , Vol. 25 , 1947, pp. 261-278 .

[34] J. Makhoul, "Linear Prediction: A Tutorial Review", *Proc. IEEE*, Vol. 63, April 1974, pp. 561-580.

[35] R.J. McEliece and J.B. Shearer , "A Property of Euclid's Algorithm and an Application to Padé Approximation", *SIAM J. Appl. Math*, Vol. 34, 1978, pp. 611-617 .

[36] M. Morf , A. Vieira and D. T. Lee , Ladder Forms for Identification and Speech Processing , *Proc. Conf. on Decision and Control* , Dec. 7-9, 1977 .

[37] G. Nemeth and M. Zimanyi, "Polynomial Type Padé Approximants", *Math. of Comp.*, Vol. 38, 1982, pp. 553-565.

[38] S. Paszkowski, "Recurrence Relations in Padé-Hermite Approximation", *J. of Comp. and Appl. Math*, Vol. 19, 1987, pp. 99-107.

[39] J. Rissanen , "Solution of Linear Equations with Hankel and Toeplitz Matrices ", *Numer. Math* , Vol. 22, 1974 , pp. 361-366.

[40] J. Rissanen, Recursive Evaluation of Padé Approximants for Matrix Sequences, *IBM J. Res. Develop.*, 1972, pp. 401-406.

[41] J. Rissanen, "Algorithms for Triangular Decomposition of Block Hankel and Toeplitz Matrices with Application to Factoring Positive Matrix Polynomials", *Math. Comp.*, Vol. 27, 1973, pp. 147-154.

[42] H. Rutishauser, "Der Quotient-Differenzen-Algorithm", *Z. Angew. Math. Phys.*, Vol. 5, 1954, pp. 233-251.

[43] Y. Starkand, "Explicit Formulas for Matrix-valued Padé Approximants", *J. of Comp. and Appl. Math*, Vol. 5, No. 1, 1979, pp. 63-65.

[44] Y. Sugiyama , "An Algorithm for Solving Discrete-time Wiener-Hopf Equations based on Euclid's Algorithm ", *IEEE Trans. Inform. Theory* , Vol. 32 , 1986, pp. 394-409 .

[45] W.F. Trench , "An Algorithm for the Inversion of Finite Hankel Matrices", *SIAM J. Appl. Math* , Vol. 13 , 1965, pp. 1102-1107 .

[46] H. Wakita, "Direct Estimation of the Vocal Tract Shape by Inverse Filtering of Acoustic Speech Waveforms", *IEEE Trans. Audio and Electroacoustics*, Vol. AU-21, No. 5, Oct 1973, pp. 417-427.

[47] R.A. Wiggens and E.A. Robinson , "Recursive Solution to the Multichannel Filtering Problem ", *J. Geophys. Res.* , Vol. 70, 1965 , pp. 1885-1891 .

[48] P. Wynn, "The Rational Approximation of Functions which are formally defined by a power series expansion", *Math. Comp.*, Vol. 14, 1960, pp. 147-186.

[49] S. Zohar , "Toeplitz Matrix Inversion: The Algorithm of W.F. Trench ", *J. ACM* , Vol. 16 , 1969, pp. 592-601 .